

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:32:03 ; Search time 12.5 Seconds
(without alignments)
184.688 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GAQFSKTAAKGEAAERPGEAAVA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	115	100.0	309	2	A39773
2	115	100.0	332	2	A38873
3	111	96.5	309	2	A39169
4	111	96.5	335	2	S08341
5	104	90.4	281	2	A41400
6	50	43.5	399	2	AC3089
7	50	43.5	399	2	G98197
8	48	41.7	261	2	AH0089
9	48	41.7	1058	2	T50496
10	47	40.9	352	2	T36719
11	46.5	40.4	1293	2	T01512
12	46	40.0	250	2	T45028
13	46	40.0	317	2	A87479
14	46	40.0	416	1	W2MLDP
15	45	39.1	405	2	T47595
16	45	39.1	409	2	S44346
17	45	39.1	421	2	E87277
18	45	39.1	923	2	A39596
19	45	39.1	1088	2	C83054
20	44.5	38.7	281	2	JC4558
21	44.5	38.7	335	2	F84411
22	44.5	38.7	811	2	T36581
23	44	38.3	194	2	A70609
24	44	38.3	287	2	A54601
25	44	38.3	346	2	A55111
26	44	38.3	346	2	JC7638
27	44	38.3	432	2	D86937
28	44	38.3	486	2	C75154
29	44	38.3	503	2	T35053

30	44	38.3	515	2	C75615
31	44	38.3	1207	2	T23754
32	43.5	37.8	661	2	C81822
33	43.5	37.8	661	2	D81055
34	43	37.4	110	1	R6BYP3
35	43	37.4	118	2	S27476
36	43	37.4	154	2	T04159
37	43	37.4	327	2	JC4385
38	43	37.4	340	1	A2HU
39	43	37.4	391	2	JQ1192
40	43	37.4	391	2	AC0691
41	43	37.4	401	2	AE3583
42	43	37.4	450	1	A43733
43	43	37.4	574	2	G71336
44	43	37.4	562	2	A45155
45	43	37.4	584	2	T40319

ALIGNMENTS

RESULT 1

A39773
myristylated alanine-rich protein kinase C substrate, macrophage - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-Feb-1992 #sequence revision 21-Feb-1992 #text_change 18-Feb-2000
C:Accession: A39773; S29267; S05681
R:Erusalimsky, J.D.; Brooks, S.F.; Herget, T.; Morris, C.; Rozengurt, E.
J. Biol. Chem. 266, 7073-7080, 1991
A:Title: Molecular cloning and characterization of the acidic 80-kDa protein kinase C sub-
A:Reference number: A39773; MUID:91201362; PMID:1707878
A:Accession: A39773
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-309 <ERU1>
A:Cross-references: GB:M59859; NID:g205068
A:Note: this translation is not annotated in GenBank entry RATKINC, release 113.0
R:Herget, T.; Brooks, S.F.; Broad, S.; Rozengurt, E.
Eur. J. Biochem. 209, 7-14, 1992
A:Title: Relationship between the major protein kinase C substrates acidic 80-kDa protein
or equivalent genes in different species.
A:Reference number: S29267; MUID:93011168; PMID:1396720
A:Accession: S29267
A:Molecule type: mRNA
A:Residues: 182,'E',184-301 <HER>
R:Erusalimsky, J.D.; Morris, C.; Perks, K.; Brown, R.; Brooks, S.; Rozengurt, E.
FEBS Lett. 255, 149-153, 1989
A:Title: Internal amino acid sequence analysis of the 80 kDa protein kinase C substrate f
A:Reference number: S05681; MUID:90005952; PMID:2676596
A:Accession: S05681
A:Molecule type: protein
A:Residues: 12-28, GX, 31,189-215,'E',217,252-275,'X',299-309 <ERU2>
A:Note: the amino end of the mature protein is blocked
C:Superfamily: neurofilament triplet H protein
C:Keywords: blocked amino end; lipoprotein; myristylation; phosphoprotein
F/2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted

Query Match 100.0%; Score 115; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 4.1e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAAKGEAAERPGEAAVA 24

Db 2 GAQFSKTAAKGEAAERPGEAAVA 25

RESULT 2

A38873
myristylated alanine-rich protein kinase C substrate - human
N:Alternate names: acidic calmodulin-binding 80K protein; MARCKS
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text_change 20-Jun-2000
C:Accession: A38873; A42977; A40758; S29269

R;Shimizu, N.
submitted to DDBJ, September 1991
A:Reference number: A38873
A:Accession: A38873
A:Molecule type: mRNA
A:Residues: 1-332 <SHI>
A:Cross-references: GB:D10522; GB:D90498; NID:g219893; PIDN:BAA01392.1; PID:g219894
R;Sakai, K.; Hirai, M.; Kudo, J.; Minoshima, S.; Shimizu, N.
Genomics 14, 175-178, 1992
A:Title: Molecular cloning and chromosomal mapping of a cDNA encoding human 80K-L protein
A:Reference number: A42977; MUID:93052291; PMID:1427823
A:Accession: A42977
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-13, 'G' 15-332 <SAK>
A:Cross-references: GB:D90498
A:Experimental source: squamous carcinoma cells A431
A:Note: sequence extracted from NCBI backbone (NCBI:118653)
R;Harlan, D.M.; Graff, J.M.; Stumpo, D.J.; Eddy Jr., R.L.; Shows, T.B.; Boyle, J.M.; Blau, H.
J. Biol. Chem. 266, 14399-14405, 1991
A:Title: The human myristoylated alanine-rich C kinase substrate (MARCKS) gene (MACS). A
A:Reference number: A40758; MUID:913117795; PMID:1860846
A:Accession: A40758
A:Molecule type: mRNA
A:Residues: 1-83, 'A' 85-118, 'P' 120-233, 'W' 235-286, 'LVC' 290, 'RRGSGRGGARRSLNQ', 309-311
A:Cross-references: GB:M68956
A:Note: the authors translated the codon GGC for residue 53 as Arg
R;Herget, T.; Brooks, S.F.; Broad, S.; Rozengurt, E.
Eur. J. Biochem. 209, 7-14, 1992
A:Title: Relationship between the major protein kinase C substrates acidic 80-kDa protein or equivalent genes in different species.
A:Reference number: S29267; MUID:93011168; PMID:1396720
A:Accession: S29269
A:Molecule type: mRNA
A:Residues: 183-223, 'A' 225-234, 'E' 236-322 <HER>
C:Comment: This protein is a major cellular substrate for protein kinase C and plays a role in calcium signaling.
C:Genetics:
A:Gene: GDB:MACS
A:Cross-references: GDB:118835; OMIM:177061
A:Map position: 6q22.2-6q22.2
C:Superfamily: neurofilament triplet H protein
C:Keywords: actin binding; blocked amino end; calmodulin binding; lipoprotein; myristylation; modified site: myristoylated amino end (Gly) (in mature form) #status predicted
F:2/Modified site: myristoylated amino end (Gly) (covalent) (by protein kinase C) #status predicted
P:159,163,167,170/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 100.0%; Score 115; DB 2; Length 332;
Best Local Similarity 100.0%; Pred. No. 4.3e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAARPGGAAVA 24
|||||
DB 2 GAQFSKTAAGGAAARPGGAAVA 25
|||||

RESULT 3
A39169
myristoylated alanine-rich protein kinase C substrate, macrophage - mouse
N:Alternate names: MARCKS; myristoylated alanine-rich C kinase substrate
C:Species: Mus musculus (house mouse)
C:Date: 24-Jan-1992 #sequence revision 24-Jan-1992 #text_change 18-Feb-2000
A:Accession: A39169; S16519; S29268; S63977; S74153
R;Seykora, J.T.; Ravetch, J.V.; Aderem, A.
Proc. Natl. Acad. Sci. U.S.A. 88, 2505-2509, 1991
A:Title: Cloning and molecular characterization of the murine macrophage "68-kDa" protein
A:Reference number: A39169; MUID:91172836; PMID:2006186
A:Accession: A39169
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-309 <SEY>
A:Cross-references: GB:M60474; NID:g199026; PIDN:AAA39491.1; PID:g199027
R;Brooks, S.F.; Herget, T.; Erusalimsky, J.D.; Rozengurt, E.
EMBO J. 10, 2497-2505, 1991

A:Title: Protein kinase C activation potentially down-regulates the expression of its major substrate, acidic 80-kDa protein
A:Reference number: S16519; MUID:91330872; PMID:1868832
A:Accession: S16519
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-95, 'T' 97, 'T' 99-309 <EMB>
R;Herget, T.; Brooks, S.F.; Broad, S.; Rozengurt, E.
Eur. J. Biochem. 209, 7-14, 1992
A:Title: Relationship between the major protein kinase C substrates acidic 80-kDa protein or equivalent genes in different species.
A:Reference number: S29267; MUID:93011168; PMID:1396720
A:Accession: S29268
A:Molecule type: mRNA
A:Residues: 182-301 <HER>
R;Herget, T.; Oehrlein, S.A.; Pappin, D.J.C.; Rozengurt, E.; Parker, P.J.
Eur. J. Biochem. 233, 448-457, 1995
A:Title: The myristoylated alanine-rich C-kinase substrate (MARCKS) is sequentially phosphorylated in different species.
A:Reference number: S63977; MUID:96067684; PMID:7588787
A:Accession: S63977
A:Molecule type: protein
A:Residues: 145-151, 'X' 153-155, 'X' 157-162, 'X' 164-169 <HEW>
R;Schoenwasser, D.C.; Palmer, R.H.; Herget, T.; Parker, P.J.
FEBS Lett. 395, 1-5, 1996
A:Title: p42 MAPK phosphorylates 80 kDa MARCKS at Ser-113.
A:Reference number: S74153; MUID:97002278; PMID:8849678
A:Accession: S74153
A:Molecule type: protein
A:Residues: 102-130 <SCH>
C:Superfamily: neurofilament triplet H protein
C:Keywords: actin binding; blocked amino end; calmodulin binding; lipoprotein; myristylation; modified site: myristoylated amino end (Gly) (in mature form) #status predicted
F:152,156,160,163/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 96.5%; Score 111; DB 2; Length 309;
Best Local Similarity 95.8%; Pred. No. 1.5e-08;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAARPGGAAVA 24
|||||
DB 2 GAQFSKTAAGGAAARPGGAAVA 25
|||||

RESULT 4
S08341
myristoylated alanine-rich protein kinase C substrate - bovine
N:Alternate names: acidic calmodulin-binding 81K protein (ACAMP-81); MARCKS
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-Dec-1990 #sequence revision 31-Dec-1990 #text_change 18-Feb-2000
A:Accession: S08341; A32904; S29270; A46098; PS0338
R;Stumpo, D.J.; Graff, J.M.; Albert, K.A.; Greengard, P.; Blackshear, P.J.
Nucleic Acids Res. 17, 3987-3988, 1989
A:Title: Nucleotide sequence of a cDNA for the bovine myristoylated alanine-rich C kinase
A:Reference number: S08341; MUID:89282412; PMID:2734111
A:Accession: S08341
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-335 <STU>
A:Cross-references: EMBL:M24638; NID:g163339; PIDN:AAA30635.1; PID:g163340
R;Stumpo, D.J.; Graff, J.M.; Albert, K.A.; Greengard, P.; Blackshear, P.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 4012-4016, 1989
A:Title: Molecular cloning, characterization, and expression of a cDNA encoding the "80-kDa" protein
A:Reference number: A32904; MUID:89264553; PMID:2726763
A:Accession: A32904
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-98, 'Q' 100-335 <ST2>
A:Cross-references: GB:M24638; GB:M23738
R;Herget, T.; Brooks, S.F.; Broad, S.; Rozengurt, E.
Eur. J. Biochem. 209, 7-14, 1992
A:Title: Relationship between the major protein kinase C substrates acidic 80-kDa protein or equivalent genes in different species.
A:Reference number: S29267; MUID:93011168; PMID:1396720
A:Accession: S29270

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, A.; Sterner, E.W.

A;Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AC3089

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-399 <KUR>

A;Cross-references: GB:AE008689; PIDN:AAL45129.1; PID:g17742801; GSPDB:GN00187

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: Atu4335

A;Map position: linear chromosome

	Query Match	Best Local Similarity	43.5%;	Score 50;	DB 2;	Length 399;
	Matches 11;	Conservative 4;	Mismatches 7;	Indels 0;	Gaps 0;	

Qy 3 QFSKTAAKGEAAARPGGAAVA 24
|||:|||||:|:
Db 367 QFLQAARGEAASARETDTAA 388

RESULT 7

G98197

hypothetical protein AGR_L1057 [imported] - *Agrobacterium tumefaciens* (strain C58, Cerevisia)

C;Species: *Agrobacterium tumefaciens*

C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002

C;Accession: G98197

R;Goodner, B.; Hinkele, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens*

A;Reference number: A97359; MUID:21608551; PMID:11743194

A;Accession: G98197

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-399 <KUR>

A;Cross-references: GB:AE007870; PIDN:AAK89105.1; PID:g15158909; GSPDB:GN00170

C;Genetics:

A;Gene: AGR_L1057

A;Map position: linear chromosome

	Query Match	Best Local Similarity	43.5%;	Score 50;	DB 2;	Length 399;
	Matches 11;	Conservative 4;	Mismatches 7;	Indels 0;	Gaps 0;	

Qy 3 QFSKTAAKGEAAARPGGAAVA 24
|||:|||||:|:
Db 367 QFLQAARGEAASARETDTAA 388

RESULT 8

AH0089

probable flagellar basal-body rod protein flgG [imported] - *Yersinia pestis* (strain CO92)

C;Species: *Yersinia pestis*

C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001

C;Accession: AH0089

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; de novo-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Fildes, I.; M. Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, E.; Nature 413, 523-527, 2001

A;Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.

A;Reference number: AB0001; MUID:21470413; PMID:11586360

A;Accession: AH0089

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-261 <KUR>

A;Cross-references: GB:AL590842; PIDN:CAC89579.1; PID:g15978811; GSPDB:GN00175

C;Genetics:

A;Gene: flgG

C;Superfamily: rod protein flgF

```

Query Match          41.7%; Score 48; DB 2; Length 261;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GAQFSKTAAGGAAARPGGAA 22
   |||:|||||
Db 189 GNLYRETAAGGAEVGVGEEA 210

RESULT 9
T50496
hypothetical protein T22D6.50 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: T50496
R:Bevan, M.; Terry, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.; De
  ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
  submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25101
A:Accession: T50496
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1058 <BEV>
A:Cross-references: EMBL:AL357612
A:Experimental source: cultivar Columbia; BAC clone T22D6
C:Genetics:
A:Map position: 5
A:Introns: 60/3; 195/3; 222/3; 448/3; 526/3; 555/2; 591/2; 616/3; 662/2; 715/3; 7
A:Note: T22D6.50

Query Match          41.7%; Score 48; DB 2; Length 1058;
Best Local Similarity 52.6%; Pred. No. 56;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 5 SKTAAGGAAARPGGAAV 23
   |||:|||||
Db 532 SKVSSSEAAADKPSGA 550

RESULT 10
T36719
probable integral membrane protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T36719
R:Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
  submitted to the EMBL Data Library, June 1999
A:Reference number: Z21612
A:Accession: T36719
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-352 <MUR>
A:Cross-references: EMBL:AL079308; PIDN: CAB45217.1; GSPDB: GN000070; SCOEDB: SCH69.20c
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB: SCH69.20c

Query Match          40.9%; Score 47; DB 2; Length 352;
Best Local Similarity 43.5%; Pred. No. 29;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 2 AQFSKTAAGGAAARPGGAAV 24
   |||:|||||
Db 114 APLSRVARRQARRKPGGA 136

RESULT 11
T01512
hypothetical protein T10M13.12 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 24-Mar-1999
C:Accession: T01512

```

```

R:Johnson, A.F.; de la Bastide, M.; Lodhi, M.; Hoffman, J.; Hasegawa, A.; Gnoj, L.; Gott
  Martienssen, R.; McCombie, W.
  submitted to the EMBL Data Library, May 1997
A:Description: The sequence of the Arabidopsis thaliana T10M13 BAC.
A:Reference number: Z14346
A:Accession: T01512
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1293 <JOH>
A:Cross-references: EMBL:AF001308; NID: g2104523; PID: g3912929
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4S
A:Introns: 29/3; 59/3; 94/3; 121/1; 144/2; 159/2; 1120/2; 1192/3
A:Note: T10M13.12

Query Match          40.4%; Score 46.5; DB 2; Length 1293;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

Qy 1 GAQ---FSKTAAGGAAARPGGAAV 23
   |||:|||||
Db 905 GAQSLVVEKSLAKKEAAADPSNAAM 930

RESULT 12
T45028
hypothetical protein Y39B6.b [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:Accession: T45028
R:Wilson, R.; Ainscough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burton,
  raser, A.; Fulton, L.; Gardner, A.; Green, P.; Hawkins, T.; Hallier, L.; Jier, M.; Johns
  B.; O'Callaghan, M.; Parsons, J.; Percy, C.; Rifken, L.; Roopra, A.; Saunders, D.
  Nature 368, 32-38, 1994
A:Authors: Showkeen, R.; Sims, M.; Smaldon, N.; Smith, A.; Smith, M.; Sonhammer, E.; S
  tock, L.; Wilkinson-Sproat, J.; Wohlman, P.
A>Title: 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.
A:Reference number: S43531; MUID: 94150718; PMID: 7906398
A:Accession: T45028
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-250 <WIL>
A:Cross-references: EMBL:AL132896; NID: g6434440; PIDN: CAB60907.1; PID: g6434442
A:Experimental source: clone Y39B6B
C:Genetics:
A:Map position: 3
A:Introns: 66/3; 129/2; 219/2
A:Note: Y39B6B.b

Query Match          40.0%; Score 46; DB 2; Length 250;
Best Local Similarity 47.4%; Pred. No. 30;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 3 QFSKTAAGGAAARPGGAA 21
   |||:|||||
Db 199 RFTQAQAKAAEAAPAKS 217

RESULT 13
A87479
heptosyltransferase family protein [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: A87479
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
  B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
  n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
  Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A>Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID: 21173698; PMID: 11259647
A:Accession: A87479
A:Status: preliminary

```


Search completed: October 7, 2004, 17:36:01
Job time : 16.5 secs

A:Molecule type: DNA
A:Residues: 1-317 <STO>
A:Cross-references: GB:AE005673; NID:gl3423295; PIDN:AAK23829.1; GSFDB:GNO0148
C:Genetics:
A:Gene: CCL854

Query Match 40.0%; Score 46; DB 2; Length 317;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 12; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
QY 1 GAQFSKTAAGGAAERPGGAAYA 24
DB 27 GARISWCAAKGETVYTGPLKAVVA 50

RESULT 14
W2WLDP
E2 protein - deer papillomavirus
C:Species: deer papillomavirus
A:Note: host Odocoileus virginianus (American white-tailed deer)
C>Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 24-Feb-1994
C:Accession: A03673
R:Groff, D.E.; Lancaster, W.D.
J Virol. 56; 85-91, 1985
A:Title: Molecular cloning and nucleotide sequence of deer papillomavirus.
A:Reference number: A93013; MUID:85293253; PMID:2993669
A:Accession: A03673
A:Molecule type: DNA
A:Residues: 1-416 <GRO>
C:Superfamily: papillomavirus E2 protein
C:Keywords: early protein

Query Match 40.0%; Score 46; DB 1; Length 416;
Best Local Similarity 60.0%; Pred. No. 47;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 9 AKGEAAERPGGAAY 23
DB 368 AVGEQGSERPGDATV 382

RESULT 15
T47595
RING finger protein T12E18.50 - Arabidopsis thaliana
N:Alternate names: protein T12E18.50
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 05-May-2000
C:Accession: T47595
R:Blocker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat M.Mewes,
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24469
A:Accession: T47595
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-405 <BLO>
A:Cross-references: EMBL:AL132971
A:Experimental source: cultivar Columbia; BAC clone T12E18
C:Genetics:
A:Gene: T12E18.50
A:Map position: 3
A:Introns: 130/2; 208/3; 257/2; 286/3; 335/3; 385/3
C:Superfamily: Arabidopsis thaliana RING finger protein T12E18.50; RING finger homology
F104-148/Domain: RING finger homology <RRN>

Query Match 39.1%; Score 45; DB 2; Length 405;
Best Local Similarity 37.5%; Pred. No. 64;
Matches 9; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
QY 1 GAQFSKTAAGGAAERPGGAAYA 24
DB 30 GCPFSKAARPPDDASARKQETTAS 53

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:32:03 ; Search time 8 seconds
(without alignments)
156,210 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GQFSKTAAGGAERPEAAVA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	115	100.0	308	1 MACS_RAT	P30009 rattus norv
2	115	100.0	331	1 MACS_HUMAN	P29966 homo sapien
3	111	96.5	308	1 MACS_MOUSE	P26645 mus musculu
4	111	96.5	331	1 MACS_BOVIN	P12624 bos taurus
5	104	90.4	280	1 MACS_CHICK	P16527 gallus gall
6	67	58.3	85	1 MRXD_RAT	P20468 rattus norv
7	46	40.0	382	1 OXAA_PROMA	Q7V800 prochloroco
8	46	40.0	416	1 VE2_FAPVD	P03123 deer papill
9	45	39.1	409	1 R23B_HUMAN	P54727 homo sapien
10	45	39.1	923	1 PRGR_MOUSE	Q00175 mus musculu
11	45	39.1	1587	1 TOP2_PENCH	Q9V868 penicillium
12	44	38.3	291	1 USF2_RAT	Q63655 rattus norv
13	44	38.3	346	1 USF2_HUMAN	Q15853 homo sapien
14	44	38.3	346	1 USF2_MOUSE	Q64705 mus musculu
15	43	37.4	110	1 RLA3_SCHPO	P17477 schizosacch
16	43	37.4	118	1 YNIF_AZOB	P25316 azospirillu
17	43	37.4	246	1 BRM1_MOUSE	Q99n20 mus musculu
18	43	37.4	306	1 TEPI_MOUSE	O54819 mus musculu
19	43	37.4	326	1 PDL1_MOUSE	O70400 mus musculu
20	43	37.4	327	1 PDL1_RAT	P25294 rattus norv
21	43	37.4	391	1 MANA_SALTY	P52081 salmoneilla
22	43	37.4	421	1 SP39_BRUAB	O06875 bruceella ab
23	43	37.4	421	1 SP39_BRUME	Q8fvc5 bruceella me
24	43	37.4	421	1 SP39_BRUSU	P14590 lemur catta
25	43	37.4	450	1 INVO_LEMCA	O95466 homo sapien
26	43	37.4	463	1 FMNL_HUMAN	O83349 treponema p
27	43	37.4	574	1 GLYA_TREPA	O05049 xenopus lae
28	43	37.4	662	1 MUC1_XENLA	O977x6 schizosacch
29	43	37.4	684	1 ISH1_SCHPO	P91327 drosophila
30	43	37.4	1013	1 A60D_DROME	Q9pfd6 xylella fas
31	43	37.4	1322	1 PUR4_XYLF	Q05152 cryptotolag
32	43	37.4	2339	1 CCAB_RABIT	P28283 herpes simp
33	42	36.5	261	1 RL1_HSV2H	

34	42	36.5	262	1 FLGG_CAUCR	Q06172 caulobacter
35	42	36.5	274	1 BPAL_STRAU	P33912 streptomyce
36	42	36.5	351	1 CSP_PLAKU	P04922 plasmodium
37	42	36.5	442	1 CY94_DICDI	P54639 dictyosteli
38	42	36.5	495	1 AB31_RHLRE	Q8s339 chlamydomon
39	42	36.5	627	1 MUTL_RHILO	Q98312 rhizobium l
40	42	36.5	645	1 DNAK_RHOBA	Q7um31 rhodopirell
41	42	36.5	733	1 HIC1_HUMAN	Q14526 homo sapien
42	42	36.5	788	1 NASP_HUMAN	P49321 homo sapien
43	42	36.5	1159	1 SOR2_HUMAN	Q9gpq0 homo sapien
44	42	36.5	1198	1 HCN4_RAT	Q9jka7 rattus norv
45	41	35.7	106	1 RLA3_YEAST	P10622 saccharomyc

ALIGNMENTS

RESULT 1

ID	MACS_RAT	STANDARD;	PRT;	308 AA.
AC	P30009;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Myristoylated alanine-rich C-kinase substrate (MARCKS).			
GN	MARCKS OR MACS.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=91201362; PubMed=1707878;			
RA	Brusallinsky J.D., Brooks S.P., Herget T., Morris C., Rozengurt E.;			
RT	"Molecular cloning and characterization of the acidic 80-kDa protein			
RT	kinase C substrate from rat brain. Identification as a			
RT	glycoprotein.";			
RL	J. Biol. Chem. 266:7073-7080(1991).			
RN	[2]			
RP	PHOSPHORYLATION SITES.			
RC	TISSUE=Brain;			
RX	MEDLINE=93135774; PubMed=8422248;			
RA	Heemskerk F.M., Chen H.C., Huang F.L.;			
RT	"Protein kinase C phosphorylates Ser-152, Ser-156 and Ser-163 but not			
RT	Ser-160 of MARCKS in rat brain.";			
RL	Biochem. Biophys. Res. Commun. 190:236-241(1993).			
CC	-!- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR			
CC	PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND			
CC	SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.			
CC	-!- PTM: PHOSPHORYLATION BY PKC DISPLACES MARCKS FROM THE MEMBRANE. IT			
CC	ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.			
CC	-!- SIMILARITY: Belongs to the MARCKS family.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M59859; -; NOT_ANNOTATED_CDS.

PIR; A39773; A39773.

InterPro; IPR002101; MARCKS.

Pfam; PF02063; MARCKS; 1.

PRINTS; PR00963; MARCKS.

PROSITE; PS00826; MARCKS 1; 1.

PROSITE; PS00827; MARCKS 2; 1.

Phosphorylation; Myristate; Calmodulin-binding; Actin-binding;

Membrane; Lipoprotein.

INIT MET 0 0 BY SIMILARITY.

LIPID 1 1 N-myristoyl glycine (By similarity).

```

FT DOMAIN 144 168 CALMODULIN-BINDING (PSD)
FT MOD_RSS 151 151 PHOSPHORYLATION (BY PKC)
FT MOD_RSS 155 155 PHOSPHORYLATION (BY PKC)
FT MOD_RSS 162 162 PHOSPHORYLATION (BY PKC)
SQ SEQUENCE 308 AA; 29663 MW; 59E50228BB1B75D2 CRC64;

Query Match 100.0%; Score 115; DB 1; Length 308;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GAQFSKTAAGGAAERPGAAVA 24
| | | | | | | | | | | | | | | | | |
Db 1 GAQFSKTAAGGAAERPGAAVA 24
| | | | | | | | | | | | | | | | | |

RESULT 2
MACS_HUMAN STANDARD; PRT; 331 AA.
ID MACS_HUMAN STANDARD; PRT; 331 AA.
AC P29966;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myristoylated alanine-rich C-kinase substrate (MARCKS) (Protein kinase
DE C substrate, 80 kDa protein, light chain) (PKCSL) (80K-L protein).
GN MARCKS OR MACS OR PRKCSL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91317795; PubMed=1860846;
RA Harlan D.M., Graff J.M., Stumpo D.J., Eddy R.L. Jr., Shows T.B.,
RA Boyle J.M., Blackshear P.J.;
RT "The human myristoylated alanine-rich C kinase substrate (MARCKS)
RT gene (MACS). Analysis of its gene product, promoter, and chromosomal
RT localization."
RL J. Biol. Chem. 266:14399-14405(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91052291; PubMed=1427823;
RA Sakai K., Hirai M., Kudo J., Minoshima S., Shimizu N.;
RT "Molecular cloning and chromosomal mapping of a cDNA encoding human
RT 80K-L protein: major substrate for protein kinase C."
RL Genomics 14:175-178(1992).
CC -!- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR
CC PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND
CC SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.
CC -!- PTM: PHOSPHORYLATION BY PKC DISPLACES MARCKS FROM THE MEMBRANE. IT
CC ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.
CC -!- SIMILARITY: Belongs to the MARCKS family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M68956; AAA59555.1; -
CC DR EMBL; M68955; AAA59554.1; -
CC DR EMBL; D10522; BAA01392.1; -
CC DR PIR; A38873; A38873.
CC DR Genew; HGNC:6759; MARCKS.
CC DR MIM; 177061; -
CC DR GO; GO:0035629; C:actin cytoskeleton; TAS.
CC DR GO; GO:0003780; F:actin cross-linking activity; TAS.
CC DR GO; GO:0005516; F:calmodulin binding; TAS.
CC DR InterPro; IPR002101; MARCKS.
CC DR Pfam; PF02063; MARCKS; 1.
CC DR PRINTS; PR00963; MARCKS.
CC DR PROSITE; PS00826; MARCKS_1; 1.

```

```

DR PROSITE; PS00827; MARCKS_2; 1.
KW Phosphorylation; Myristate; Calmodulin-binding; Actin-binding;
KW Membrane; Lipoprotein.
FT INIT_MET 0 0 BY SIMILARITY.
FT LIPID 1 1 N-myristoyl glycine (By similarity).
FT DOMAIN 151 175 CALMODULIN-BINDING (PSD).
FT MOD_RSS 158 158 PHOSPHORYLATION (BY PKC).
FT MOD_RSS 162 162 PHOSPHORYLATION (BY PKC).
FT MOD_RSS 166 166 PHOSPHORYLATION (BY PKC).
FT MOD_RSS 169 169 PHOSPHORYLATION (BY PKC).
FT CONFLICT 83 83 S -> A (IN REF. 1).
FT CONFLICT 118 118 A -> P (IN REF. 1).
FT CONFLICT 233 233 P -> S (IN REF. 1).
FT CONFLICT 286 307 PGAPPEQFAAEFAAEFAAAASS -> LVCPRRGSGPRGAR
FT GRSILNO (IN REF. 1).
SQ SEQUENCE 331 AA; 31413 MW; BCC837D586581774 CRC64;

Query Match 100.0%; Score 115; DB 1; Length 331;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GAQFSKTAAGGAAERPGAAVA 24
| | | | | | | | | | | | | | | | | |
Db 1 GAQFSKTAAGGAAERPGAAVA 24
| | | | | | | | | | | | | | | | | |

RESULT 3
MACS_MOUSE STANDARD; PRT; 308 AA.
ID MACS_MOUSE STANDARD; PRT; 308 AA.
AC P26645;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Myristoylated alanine-rich C-kinase substrate (MARCKS).
GN MARCKS OR MACS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Macrophage;
RX MEDLINE=91172836; PubMed=2006186;
RA Seykora J.T., Ravetch J.V., Aderem A.;
RT "Cloning and molecular characterization of the murine macrophage '68-
RT kDa' protein kinase C substrate and its regulation by bacterial
RT lipopolysaccharide."
RL Proc. Natl. Acad. Sci. U.S.A. 88:2505-2509(1991).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX TISSUE=Fibroblast;
RX MEDLINE=91330872; PubMed=1868832;
RA Brooks S.F., Herget T., Erusalimsky J.D., Rozengurt E.;
RT "Protein kinase C activation potentially down-regulates the expression
RT of its major substrate, 80K, in Swiss 3T3 cells."
RL EMBO J. 10:2497-2505(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Olfactory epithelium;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udelsin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

```

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmitt J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RL human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [4]
 RP PARTIAL SEQUENCE.
 RC STRAIN=Swiss; TISSUE=Fibroblast;
 RX MEDLINE=90346162; PubMed=2384168;
 RA Brooks S.F., Brusilinsky J.D., Totty N.F., Rozengurt E.;
 RT "Purification and internal amino acid sequence of the 80 kDa protein
 RT kinase C substrate from Swiss 3T3 fibroblasts. Homology with
 RT substrates from brain.";
 RL FEBS Lett. 268:291-295 (1990).
 CC -!- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR
 CC PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND
 CC SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.
 CC -!- TISSUE SPECIFICITY: BRAIN, SPLEEN, LESS IN KIDNEY AND HEART, AND
 CC VERY LOW LEVELS IN LIVER.
 CC -!- INDUCTION: BY lipopolysaccharides (LPS).
 CC -!- PTM: PHOSPHORYLATION BY PKC DISPLACES MARCKS FROM THE MEMBRANE. IT
 CC ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.
 CC -!- SIMILARITY: Belongs to the MARCKS family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M60474; AAA39491.1; -;
 DR EMBL; BC046601; AAA46601.1; -;
 DR PIR; A39169; A39169.
 DR MGD; MGI:96907; Marcks.
 DR InterPro; IPR002101; MARCKS.
 DR Pfam; PF02063; MARCKS; 1.
 DR PRINTS; PR00963; MARCKS.
 DR PROSITE; PS00826; MARCKS_1; 1.
 DR PROSITE; PS00827; MARCKS_2; 1.
 KW Phosphorylation; Myristate; Calmodulin-binding; Actin-binding;
 KW Membrane; Lipoprotein.
 FT INIT MET 0
 FT LIPID 1
 FT DOMAIN 144 168
 FT MOD_RES 151 151
 FT MOD_RES 155 155
 FT MOD_RES 159 159
 FT MOD_RES 162 162
 FT MOD_RES 162 162
 FT CONFLICT 95 97
 FT SEQUENCE 308 AA; 29530 MW; FB5313B913701C5C CRC64;
 Query Match 96.5%; Score 111; DB 1; Length 308;
 Best Local Similarity 95.8%; Pred. No. 8.8e-09;
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GAQFSKTAAGKGAARERGEAAVA 24
 Db 1 GAQFSKTAAGKGAARERGEAAVA 24
 RESULT 4
 MACS_BOVIN
 ID MACS_BOVIN STANDARD; PRT; 331 AA.
 AC P12624;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Myristoylated alanine-rich C-kinase substrate (MARCKS) (ACAMP-81).
 GN MARCKS OR MACS.

OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89282412; PubMed=2734111;
 RA Stumpo D.J., Graff J.M., Albert K.A., Greengard P., Blackshear P.J.;
 RT "Nucleotide sequence of a cDNA for the bovine myristoylated
 RT alanine-rich C kinase substrate (MARCKS).";
 RL Nucleic Acids Res. 17:3987-3988 (1989).
 RN [2]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=89264553; PubMed=2726763;
 RA Stumpo D.J., Graff J.M., Albert K.A., Greengard P., Blackshear P.J.;
 RT "Molecular cloning, characterization, and expression of a cDNA
 RT encoding the '80- to 87-kDa' myristoylated alanine-rich C kinase
 RT substrate: a major cellular substrate for protein kinase C.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4012-4016 (1989).
 RN [3]
 RP PARTIAL SEQUENCE.
 RX MEDLINE=92171958; PubMed=1540183;
 RA Mizutani A., Tokumitsu H., Hidaka H.;
 RT "Acidic calmodulin binding protein, ACAMP-81, is MARCKS protein
 RT interacting with synapsin I.";
 RL Biochem. Biophys. Res. Commun. 182:1395-1401 (1992).
 RN [4]
 RP PHOSPHORYLATION SITES.
 RX MEDLINE=89308594; PubMed=2473066;
 RA Graff J.M., Stumpo D.J., Blackshear P.J.;
 RT "Characterization of the phosphorylation sites in the chicken and
 RT bovine myristoylated alanine-rich C kinase substrate protein, a
 RT prominent cellular substrate for protein kinase C.";
 RL J. Biol. Chem. 264:11912-11919 (1989).
 RN [5]
 RP PHOSPHORYLATION SITES, AND REVISIONS.
 RC TISSUE=Brain;
 RX MEDLINE=94308052; PubMed=8034575;
 RA Taniguchi H., Manenti S., Suzuki M., Ticani K.;
 RT "Myristoylated alanine-rich C kinase substrate (MARCKS), a major
 RT protein kinase C substrate, is an in vivo substrate of
 RT proline-directed protein kinase(s). A mass spectroscopic analysis of
 RT the post-translational modifications.";
 RL J. Biol. Chem. 269:18299-18302 (1994).
 RN [6]
 RP REVERSIBLE ASSOCIATION WITH THE MEMBRANE.
 RX MEDLINE=91238951; PubMed=2034276;
 RA Thelen M., Rosen A., Nairn A.C., Aderem A.;
 RT "Regulation by phosphorylation of reversible association of a
 RT myristoylated protein kinase C substrate with the plasma membrane.";
 RL Nature 351:320-322 (1991).
 RN [7]
 RP ACTIN-FILAMENT CROSS-LINKING.
 RX MEDLINE=92220195; PubMed=1560845;
 RA Hartwig J.H., Thelen M., Rosen A., Janney P.A., Nairn A.C.,
 RA Aderem A.;
 RT "MARCKS is an actin filament crosslinking protein regulated by
 RT protein kinase C and calcium-calmodulin.";
 RL Nature 356:618-622 (1992).
 CC -!- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR
 CC PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND
 CC SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.
 CC -!- SIMILARITY: Belongs to the MARCKS family.
 CC -!- CAUTION: REF.1 AND REF.2 SEQUENCES DIFFER FROM THAT SHOWN IN
 CC POSITIONS 140 TO 150 DUE TO A FRAMESHIFT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; M24638; AAA30635.1; ALT_FRAME.
CC PIR; S08341; A08341.
CC InterPro: IPR002101; MARCKS.
CC Pfam; PF02063; MARCKS; 1.
CC PRINTS; PR00963; MARCKS.
CC PROSITE; PS00826; MARCKS_1; 1.
CC PROSITE; PS00827; MARCKS_2; 1.
CC KW Phosphorylation; Myristate; Calmodulin-binding; Actin-binding;
CC Membrane; Lipoprotein.
CC FT INIT_MET 0 0
CC FT LIPID 1 1 N-myristoyl glycine.
CC FT DOMAIN 150 174 CALMODULIN-BINDING (PSD).
CC FT MOD_RES 26 26 PHOSPHORYLATION.
CC FT MOD_RES 45 45 PHOSPHORYLATION.
CC FT MOD_RES 80 80 PHOSPHORYLATION.
CC FT MOD_RES 99 99 PHOSPHORYLATION.
CC FT MOD_RES 116 116 PHOSPHORYLATION.
CC FT MOD_RES 133 133 PHOSPHORYLATION.
CC FT MOD_RES 157 157 PHOSPHORYLATION (BY PKC).
CC FT MOD_RES 161 161 PHOSPHORYLATION (BY PKC).
CC FT MOD_RES 165 165 PHOSPHORYLATION (BY PKC).
CC FT MOD_RES 168 168 PHOSPHORYLATION (BY PKC).
CC SQ SEQUENCE 331 AA; 31450 MW; 0877BF117EDB35CA CRC64;

Query Match 96.5%; Score 111; DB 1; Length 331;
Best Local Similarity 95.8%; Pred. No. 9.4e-09;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAQFSKTAAGGAAERPGEAAVA 24
   |||||
Db 1 GAQFSKTAAGGAAERPGEAAVA 24

RESULT 5
MACS CHICK
ID MACS CHICK STANDARD; PRT; 280 AA.
AC P16527;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myristoylated alanine-rich C-kinase substrate (MARCKS).
OS Gallus gallus (Chicken).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
CC Gallus.
CC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90114197; PubMed=2608063;
RA Graff J.M., Stumpo D.J., Blackshear P.J.;
RT "Molecular cloning, sequence, and expression of a cDNA encoding the
RT chicken myristoylated alanine-rich C kinase substrate (MARCKS).";
RL Mol. Endocrinol. 3:1903-1906(1989).
CC -!- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR
CC PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND
CC SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.
CC -!- PTM: PHOSPHORYLATION BY PKC REPLACES MARCKS FROM THE MEMBRANE. IT
CC ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.
CC -!- SIMILARITY: Belongs to the MARCKS family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; M31650; AAA48946.1; -.
```

```
DR PIR; A41400; A41400.
DR InterPro: IPR002101; MARCKS.
DR Pfam; PF02063; MARCKS; 1.
DR PRINTS; PR00963; MARCKS.
DR PROSITE; PS00826; MARCKS_1; 1.
DR PROSITE; PS00827; MARCKS_2; 1.
DR KW Phosphorylation; Myristate; Calmodulin-binding; Actin-binding;
DR Membrane; Lipoprotein.
DR FT INIT_MET 0 0
DR FT LIPID 1 1 N-myristoyl glycine.
DR FT DOMAIN 116 140 CALMODULIN-BINDING (PSD).
DR FT MOD_RES 123 123 PHOSPHORYLATION (BY PKC).
DR FT MOD_RES 127 127 PHOSPHORYLATION (BY PKC).
DR FT MOD_RES 131 131 PHOSPHORYLATION (BY PKC).
DR FT MOD_RES 134 134 PHOSPHORYLATION (BY PKC).
DR SQ SEQUENCE 280 AA; 27597 MW; DF4E9DCC0B0839E CRC64;

Query Match 90.4%; Score 104; DB 1; Length 280;
Best Local Similarity 87.5%; Pred. No. 8.2e-08;
Matches 21; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAQFSKTAAGGAAERPGEAAVA 24
   |||||
Db 1 GAQFSKTAAGGAAERPGEAAVA 24

RESULT 6
MRKD RAT
ID MRKD RAT STANDARD; PRT; 85 AA.
AC P20468;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE protein kinase C substrate 80 kDa protein (Fragments).
OS Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=90005952; PubMed=2676596;
RA Erusalimsky J.D., Morris C., Perks K., Brown R., Brooks S.,
RA Rozengurt E.;
RT "Internal amino acid sequence analysis of the 80 kDa protein kinase C
RT substrate from rat brain: relationship to the 87 kDa substrate from
RT bovine brain";
RL FEBS Lett. 255:149-153(1989).
CC -!- FUNCTION: Cellular substrate for protein kinase C.
CC -!- SIMILARITY: TO 87 MYRISTOYLATED ALANINE-RICH C KINASE SUBSTRATE
CC (MARCKS).
CC Phosphorylation.
KW NON_TER 1 1
FT NON_CONS 20 21
FT NON_CONS 49 50
FT NON_CONS 73 74
FT NON_CONS 73 74
SQ SEQUENCE 85 AA; 7593 MW; 39501D788C12E268 CRC64;

Query Match 58.3%; Score 67; DB 1; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.0056;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 GEAAARFGEAAVA 24
   |||||
Db 1 GEAAARFGEAAVA 14

RESULT 7
OXAA PROMA
ID OXAA PROMA STANDARD; PRT; 382 AA.
AC Q7VB00;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
```

DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Inner membrane protein oxaA.
 GN OXAA OR PRO1302.
 OS Prochlorococcus marinus.
 OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
 CC Prochlorococcus.
 CX NCBI_TaxID=1219;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SARG / CCMF 1375 / SS120;
 RX MEDLINE=22810154; PubMed=12917486;
 RA Dufresne A., Salanoubat M., Partensky F., Artiguenave F., Axmann I.M.,
 RA Barbe V., Duprat S., Galperin M.Y., Koonin E.V., Le Gall F.,
 RA Makarova K.S., Ostrowski M., Oztas S., Robert C., Rogozin I.B.,
 RA Scanlan D.J., Tandeau de Marsac N., Weissenbach J., Wincker P.,
 RA Wolf Y.I., Hess W.R.;
 RT "Genome sequence of the cyanobacterium Prochlorococcus marinus SS120,
 RT a nearly minimal oxyphototrophic genome.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025(2003).
 CC !- FUNCTION: Required for the insertion of integral membrane proteins
 CC into the membrane. Probably plays an essential role in the
 CC integration of proteins of the respiratory chain complexes.
 CC Involved in integration of membrane proteins that insert
 CC dependently and independently of the Sec translocase complex (By
 CC similarity).
 CC !- SUBUNIT: Specifically interacts with transmembrane segments of
 CC nascent integral membrane proteins during membrane integration (By
 CC similarity).
 CC !- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (By similarity).
 CC !- SIMILARITY: Belongs to the OXA1/oxaA family. Subfamily 1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE017164; AAQ00346.1; -
 CC HAMAP; MF 01810; -; 1.
 KW Transmembrane; Inner membrane; Complete proteome.
 FT TRANSMEM 20 42 POTENTIAL.
 FT TRANSMEM 266 288 POTENTIAL.
 FT TRANSMEM 303 325 POTENTIAL.
 SQ SEQUENCE 382 AA; 41425 MW; 12101B3452EB9345 CRC64;
 Query Match 40.0%; Score 46; DB 1; Length 382;
 Best Local Similarity 39.1%; Pred. No. 24;
 Matches 9; Conservative 4; Mismatches 10; Indels 0; Gaps 0;
 Qy 1 GAQFSKTAAGKGAARPGAAV 23
 Db 210 GADLASVSADGTVTAKYFGDATV 232
 RESULT 8
 ID VE2 PAPVD STANDARD; PRT; 416 AA.
 AC P03123;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Probable regulatory protein E2.
 GN E2.
 OS Deer papillomavirus.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 CC Papillomavirus.
 CX NCBI_TaxID=10564;
 RN [1]
 RP SEQUENCE FROM N.A. PubMed=2993669;
 RX MEDLINE=85293253; PubMed=2993669;

RA Groff D.E., Lancaster W.D.;
 RT "Molecular cloning and nucleotide sequence of deer papillomavirus.";
 RL J. Virol. 56:85-91(1985).
 CC !- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
 CC IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACNNNNNGGT-3') PRESENT
 CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
 CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING ON E2RE'S POSITION
 CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
 CC BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
 CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
 CC REPLICATION.
 CC !- SUBUNIT: Binds DNA as a dimer (By similarity).
 CC !- SUBCELLULAR LOCATION: Nuclear.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M11910; AAA66844.1; -
 CC HSP; P03122; 2BOP.
 DR InterPro; IPR000427; E2_C.
 DR InterPro; IPR001866; E2_N.
 DR InterPro; IPR009021; Viral_DNA_bd.
 DR Pfam; PF00511; E2_C; 1.
 DR Pfam; PF00508; E2_N; 1.
 DR ProDom; PD000672; E2_C; 1.
 DR ProDom; PD000678; E2_N; 1.
 KW Early protein; Transcription regulation; Activator; DNA-binding;
 KW Trans-acting factor; DNA replication; Repressor; Nuclear protein.
 SQ SEQUENCE 416 AA; 46258 MW; 79064F8017368693 CRC64;
 Query Match 40.0%; Score 46; DB 1; Length 416;
 Best Local Similarity 60.0%; Pred. No. 26;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 Qy 9 AKGEAAARPGAAV 23
 Db 368 AVGEQSGRRPGDATV 382
 RESULT 9
 ID R23B HUMAN STANDARD; PRT; 409 AA.
 AC P54727; Q8WUB0;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE UV excision repair protein RAD23 homolog B (HHR23B) (XP-C repair
 DE complementing complex 58 kDa protein) (P58).
 GN RAD23B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=94222030; PubMed=8168482;
 RA Masutani C., Sugawara K., Yanagisawa J., Sonoyama T., Ui M.,
 RA Enomoto T., Takio K., Tanaka K., van der Spek P.J., Bootsma D.,
 RA Hoeijmakers J.H.J., Hanaoka F.;
 RT "Purification and cloning of a nucleotide excision repair complex
 RT involving the Xeroderma pigmentosum group C protein and a human
 RT homologue of yeast RAD23.";
 RL EMBO J. 13:1831-1843(1994).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANT VAL-249.
 RA Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W.,
 RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
 RA Schackwitz W.S., Sherwood J.K., Wittrak L.A., Nickerson D.A.;

```

RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Ramsay H.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A., AND VARIANT VAL-249.
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Scheuer C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh F.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield V.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Involved in DNA excision repair. May play a part in DNA
damage recognition and/or in altering chromatin structure to allow
access by damage-processing enzymes.
CC -!- SUBUNIT: Heterodimer of a 125 kDa subunit (p125) and of a
58 kDa subunit (p58). Interacts with MJD.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- DOMAIN: The ubiquitin-like domain mediates interaction with MJD.
CC -!- SIMILARITY: Contains 1 ubiquitin-like domain.
CC -!- SIMILARITY: Contains 2 UBA domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D12090; BAA04652.1; -
CC EMBL; AL137852; CAD13275.1; -
CC EMBL; AY165178; RAN47194.1; -
CC EMBL; BC020973; AAH20973.1; -
CC HSSP; P54346; S44346.
CC PIR; S4346; S44346.
CC HSSP; P54725; LDV0.
CC Genew; HGNC:9813; RAD23B.
CC MIM; 600062; -
CC GO; GO:0005634; C:nucleus; TAS.
CC GO; GO:0003697; F:single-stranded DNA binding; TAS.
CC GO; GO:0006289; P:nucleotide-excision repair; TAS.
CC InterPro; IPR004806; Rad23.
CC InterPro; IPR006636; STI1.
CC InterPro; IPR000449; UBA domain.
CC InterPro; IPR000626; Ubiquitin.
CC Pfam; PF00627; UBA; 2.
CC SMART; SM00240; ubiquitin; 1.
CC SMART; SM00727; STI1; 1.
CC SMART; SM00165; UBA; 2.
CC SMART; SM00213; UBQ; 1.
CC TIGRFAMs; TIGR00601; rad23; 1.
CC PROSITE; PS0030; UBA; 2.
CC PROSITE; PS00533; UBIQUITIN 2; 1.
KW DNA damage; DNA repair; Nuclear protein; Polymorphism; Repeat.
FT DOMAIN 1 79 UBIQUITIN-LIKE.
FT DOMAIN 188 228 UBA 1.

```

```

FT DOMAIN 364 404 UBA 2.
FT DOMAIN 103 106 POLY-THR.
FT DOMAIN 254 260 POLY-ALA.
FT DOMAIN 261 269 POLY-THR.
FT DOMAIN 336 348 POLY-GLY.
FT VARIANT 249 249
FT SEQUENCE 409 AA; 43171 MW; C026C78273BCB289 CRC64;
/FTID=VAR_014350.

Query Match 39.1%; Score 45; DB 1; Length 409;
Best Local Similarity 55.6%; Pred. No. 35;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 7 TAAKGEAAARPPGEAAVA 24
141 SAAQKEPAEKPAETPVA 158
Db

RESULT 10
PRGR MOUSE STANDARD; PRT; 923 AA.
AC Q00175;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Progesterone receptor (PR).
GN PGR OR NR3C3 OR PR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91299759; PubMed=2069958;
RA Schott D.R., Shyamala G., Schneider W., Parry G.;
RT "Molecular cloning, sequence analyses, and expression of
complementary DNA encoding murine progesterone receptor.";
RL Biochemistry 30:7014-7020(1991).
RN [2]
RP SEQUENCE OF 1-9 FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=95100331; PubMed=7802637;
RA Hagihara K., Wu-Peng X.S., Funabashi T., Kato J., Pfaff D.W.;
RT "Nucleic acid sequence and kinase hypersensitive sites of the 5'
region of the mouse progesterone receptor gene.";
RL Biochem. Biophys. Res. Commun. 205:1093-1101(1994).
CC -!- FUNCTION: The steroid hormones and their receptors are involved in
the regulation of eukaryotic gene expression and affect cellular
proliferation and differentiation in target tissues.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DOMAIN: Composed of three domains: a modulating N-terminal domain,
a DNA-binding domain and a C-terminal steroid-binding domain.
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M68915; AAA39971.1; -
CC EMBL; U12644; AAA66067.1; -
CC PIR; A39596; A39596.
CC HSSP; P06481; 1A28.
CC TRANSFAC; T04680; -.
CC MGD; MGI:97567; Pgr.
CC InterPro; IPR000536; Hormone rec lig.
CC InterPro; IPR000128; Progesterone receptor.
CC InterPro; IPR001723; Steroid receptor.
CC InterPro; IPR008946; Str_ncl_receptor.

```



```
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF02161; Prog_receptor; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STROHORMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR ProDom; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOLI; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR RECEPTOR; 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger; Steroid-binding.
FT DOMAIN 1 556 MODULATING, PRO-RICH.
FT DNA_BIND 557 622 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 557 577 C4-TYPE.
FT ZN_FING 593 617 C4-TYPE.
FT DOMAIN 671 923 STEROID-BINDING.
FT DOMAIN 184 188 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ SEQUENCE 923 AA; 99073 MW; 9415F1ED343BEE3F CRC64;

Query Match 39.1%; Score 45; DB 1; Length 923;
Best Local Similarity 52.4%; Pred. No. 75;
Matches 11; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 4 FSKTAAGGAAERPGGAAVA 24
DB 415 FFLAPAPQAAPSSRPGGAAVA 435

RESULT 11
TOP2_PENCH STANDARD; PRT; 1587 AA.
AC Q9Y8G8;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DNA topoisomerase II (EC 5.99.1.3).
GN TOP2.
OS Penicillium chrysogenum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
OX NCBI_TaxID=5076;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim K., Akashi T., Mizuguchi I., Ozeki M., Kanbe T., Kikuchi A.;
RT "Type II DNA topoisomerase of Penicillium chrysogenum."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Control of topological states of DNA by transient
CC breakage and subsequent rejoining of DNA strands. Topoisomerase II
CC makes double-strand breaks.
CC -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC of double-stranded DNA.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- MISCELLANEOUS: Eukaryotic topoisomerase I and II can relax both
CC negative and positive supercoils, whereas prokaryotic enzymes
CC relax only negative supercoils.
CC -!- SIMILARITY: Belongs to the type II topoisomerase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB029613; BAA82356.1; -.
CC HSSP; P06786; 1BGW.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR003957; CBPA_NFYB_topis.
DR InterPro; IPR001241; DNA_topoisol.
DR InterPro; IPR002205; DNA_topoisolv.
```

```
DR Pfam; PF00204; DNA_gyraseB; 1.
DR Pfam; PF00521; DNA_topoisolv; 1.
DR Pfam; PF02518; HATPase_C; 1.
DR PRINTS; PR00615; CCAATSUBUNTA.
DR PRINTS; PR00418; TPI2FAMILY.
DR ProDom; PD000742; DNA_topoisolv; 1.
DR SMART; SM00387; HATPase_C; 1.
DR SMART; SM00433; TOP2C; 1.
DR SMART; SM00434; TOP4C; 1.
DR PROSITE; PS00177; TOPOISOMERASE II; 1.
KW Isomerase; Topoisomerase; DNA-binding; ATP-binding; Phosphorylation;
KW Nuclear protein.
FT NP_BIND 196 201 ATP (POTENTIAL).
FT ACT_SITE 833 833 DNA CLEAVAGE (BY SIMILARITY).
SQ SEQUENCE 1587 AA; 178760 MW; E3A992F63474BD64 CRC64;

Query Match 39.1%; Score 45; DB 1; Length 1587;
Best Local Similarity 57.9%; Pred. No. 1.2e+02;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 6 KTAAGGAAERPGGAAVA 24
DB 1524 KPAAGKGAAGKSTAAAA 1542

RESULT 12
USF2_RAT STANDARD; PRT; 291 AA.
AC Q63665;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Upstream stimulatory factor 2 (Upstream transcription factor 2)
DE (Major late transcription factor 2) (Fragment).
GN USF2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=96139468; PubMed=8576131;
RA Viollet B., Lefrancois-Martinez A.-M., Henrion A., Kahn A.,
RA Raymond-Jean M., Martinez A.;
RT "Immunochemical characterization and transacting properties of
RT upstream stimulatory factor isoforms."
RL J. Biol. Chem. 271:1405-1415(1996).
CC -!- FUNCTION: Transcription factor that binds to a symmetrical DNA
CC sequence (E-boxes) (5'-CAGGTG-3') that is found in a variety of
CC viral and cellular promoters.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein. Binds DNA as an homodimer or a heterodimer
CC (USF1/USF2).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=1;
CC Comment=At least 2 isoforms are produced;
CC Name=1;
CC IsoId=Q63665-1; Sequence=Displayed;
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X90823; CAA62338.1; -.
CC HSSP; P22415; 1AN4.
DR TRANSFAC; T02115; -.
```

```
DR InterPro: IPR001092; HLH_basic.
DR Pfam: PF00010; HLH; 1.
DR SMART: SM00353; HLH; 1.
DR PROSITE: PS00888; HLH; 1.
KW DNA-binding; Nuclear protein; Transcription regulation;
KW Alternative splicing.
FT NON_TER 1
FT DNA_BIND 180 193
FT DOMAIN 194 236
FT DOMAIN 252 273
SQ SEQUENCE 291 AA; 13651 MW; A241C6E9AF6D2424 CRC64;

Query Match 38.3%; Score 44; DB 1; Length 291;
Best Local Similarity 41.7%; Pred. No. 36;
Matches 10; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy 1 GAQFSKTAAGGAAARPGAAVA 24
Db 54 GGQQAQTQGVGDGAQRGPAAS 77

RESULT 13
USF2 HUMAN STANDARD; PRT; 346 AA.
ID USF2_HUMAN STANDARD; PRT; 346 AA.
AC Q15853; Q00671; Q00709; Q05750; Q07952; Q15851; Q15852;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Upstream stimulatory factor 2 (USF2) (Upstream transcription factor 2)
DE (FOS-interacting protein) (FIP) (Major late transcription factor 2).
GN USF2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=96139468; PubMed=8576131;
RA Viollet B., Lefrancois-Martinez A.-M., Henrion A., Kahn A.,
RA Raymondjean M., Martinez A.;
RT "Immunological characterization and transacting properties of
RT upstream stimulatory factor isoforms."
RL J. Biol. Chem. 271:1405-1415(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97127588; PubMed=8954795;
RA Groenen P.M.A., Garcia E., Debeer P., Devriendt K., Frys J.-P.,
RA van de Ven W.J.M.;
RT "Structure, sequence, and chromosome 19 localization of human USF2
RT and its rearrangement in a patient with multicystic renal
RT dysplasia."
RL Genomics 38:141-148(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ussid T.B., Tohyiuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.O., Hulyk S.W.,
RA
```

Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahney J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield J.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[5]
SEQUENCE OF 46-346 FROM N.A.
RP TISSUE=B-cell;
RX MEDLINE=93082094; PubMed=1450663;
RA Sirito M., Walker S., Lin Q., Kozlowski M.T., Klein W.H., Sawadogo M.;
"Members of the USF family of helix-loop-helix proteins bind DNA as
RT homo- as well as heterodimers.";
RL Gene Expr. 2:231-240(1992).
RN [6]
SEQUENCE OF 93-346 FROM N.A.
RP MEDLINE=92271211; PubMed=1589769;
RA Blau M.A., Rutter W.J.;
RT "Interaction cloning: identification of a helix-loop-helix zipper
RT protein that interacts with c-Fos.";
RL Science 256:1014-1018(1992).
CC -!- FUNCTION: Transcription factor that binds to a symmetrical DNA
CC sequence (E-boxes) (5'-CACGTG-3') that is found in a variety of
CC viral and cellular promoters.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein. Binds DNA as an homodimer or a heterodimer
CC (USF1/USF2). In vivo, the USF1/USF2 heterodimer represents over
CC 66% of the USF binding activity whereas the USF1 and USF2A
CC homodimers represent less than 10%. The USF1/USF2 heterodimer
CC accounted for almost 15% in some cell.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Additional isoforms seem to exist;
CC Name=USF2A;
CC IsoId=Q15853-1; Sequence=Displayed;
CC Name=USF2A-delta-H;
CC IsoId=Q15853-2; Sequence=VSP_002165;
CC Name=USF2B;
CC IsoId=Q15853-3; Sequence=VSP_002164;
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X90826; CAA62341.1; -
CC EMBL; X90825; CAA62340.1; -
CC EMBL; X90824; CAA62339.1; -
CC EMBL; BC049821; AAH49821.1; -
CC EMBL; Y07661; CAA68942.1; -
CC EMBL; AD000684; AAB51179.1; -
CC EMBL; S50537; AAB24368.1; -
CC EMBL; M77476; -; NOT_ANNOTATED_CDS.
CC HSSP; P22415; 1AM4.
CC TRANSFAC; T00878; -
CC Genew; HGNC:12594; USF2.
CC MIM; 600390;
CC GO; GO:0003702; F-RNA polymerase II transcription factor acti. . .; TAS.
CC GO; GO:0003700; F:transcription factor activity; TAS.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS00888; HLH; 1.
CC

```

KW DNA-binding; Nuclear protein; Transcription regulation;
KW Alternative splicing.
FT DOMAIN 11 20 POLY-ALA.
FT DOMAIN 245 248 POLY-ARG.
FT DNA_BIND 235 248 BASIC DOMAIN.
FT DOMAIN 243 291 HELIX-LOOP-HELIX MOTIF.
FT DOMAIN 307 328 LEUCINE-ZIPPER (POTENTIAL).
FT VARSPLIC 77 143 Missing (in isoform USF2B).
FT VARSPLIC 275 282 /FTID-VSP 002164.
FT VARSPLIC 275 282 Missing (in isoform USF2A-delta-H).
FT CONFLICT 46 64 QTAVALTSVQQAFAFGHNI -> GGTSGGRSGGIQTRVQH
FT CONFLICT 93 100 GDTAGAVS -> EFHSWRRH (IN REF. 6).
FT CONFLICT 122 122 A -> V (IN REF. 6).
SQ SEQUENCE 346 AA; 36955 MW; 78CFE97AC4C10CF CRC64;

Query Match
Best Local Similarity 38.3%; Score 44; DB 1; Length 346;
Matches 10; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGEAAERPGAAVA 24
Db 109 GQQAVTQGVGDGAQRPGPAAAS 132

RESULT 14
USF2 MOUSE
ID USF2 MOUSE STANDARD; PRT; 346 AA.
AC Q64705;
DC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Upstream stimulatory factor 2 (Upstream transcription factor 2)
DE (Major late transcription factor 2).
GN USF2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv; TISSUE=Kidney, and Spleen;
RX MEDLINE=95014112; PubMed=7523363;
RA Lin Q., Luo X., Sawadogo M.;
RT "Archaic structure of the gene encoding transcription factor USF.";
RL J. Biol. Chem. 269:23894-23903 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94173694; PubMed=8127680;
RA Sirito M., Lin Q., Maity T., Sawadogo M.;
RT "Ubiquitous expression of the 43- and 44-kDa forms of transcription factor USF in mammalian cells.";
RL Nucleic Acids Res. 22:427-433 (1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=95293408; PubMed=7774954;
RA Henrion A.A., Martinez A., Mattei M.-G., Raymondjean M.;
RT "Structure, sequence, and chromosomal location of the gene for USF2 transcription factors in mouse.";
RL Genomics 25:36-43 (1995).
RN [4]
RP SEQUENCE FROM N.A.
CC -!- FUNCTION: Transcription factor that binds to a symmetrical DNA sequence (E-boxes) (5'-CACGTG-3') that is found in a variety of viral and cellular promoters.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another bHLH protein. Binds DNA as an homodimer or a heterodimer (USF1/USF2).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=USF2A;

```

RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez E.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: Plays an important role in the elongation step of
CC protein synthesis.
CC -!- SUBUNIT: P1 and P2 exist as dimers at the large ribosomal subunit.
CC -!- MISCELLANEOUS: Yeasts contain 4 individual small ribosomal A
CC proteins (RPA) which can be classified into two couples of similar
CC but not identical sequences. Each couple is distinctly related to
CC one of the two A proteins present in multicellular organisms.
CC -!- MISCELLANEOUS: Rpa3 and rpa4 are essential for cell survival,
CC whereas rpa1 and rpa2 are not.
CC -!- SIMILARITY: Belongs to the L12P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M33139; AAA35336.1; -;
CC EMBL; AL022070; CAA17793.1; -;
CC PIR; C34715; R6BYP3.
CC GeneDB.SPombe; SPBC3B9.13c; -;
CC InterPro; IPR001813; Ribosomal_60S.
CC Pfam; PF00428; 60s_ribosomal; 1.
CC Ribosomal protein; Phosphorylation; Multigene family.
SQ SEQUENCE 110 AA; 11171 MW; 1734AC9779F5A891 CRC64;

Query Match 37.4%; Score 43; DB 1; Length 110;
Best Local Similarity 54.5%; Pred. No. 20;
Matches 12; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GAQFGKTAAGGAARPGEEA 22
||| |||||
Db 73 GAGAPAAAGGAARAEQKEA 94

Q8h321 oryza sativ
Q8h452 rhizobium l
Q8h587 streptomyces
Q8h611 rhizodiporell
Q8h8x8 streptomyces
Q9vyp7 trophomyia
Q9xat4 streptomyces
Q8h7d3 pseudomonas
Q92q19 rhizobium m
Q92q19 bordetella
Q7w469 bordetella
Q9w6b1 brachydanio
Q7y415 enterobacte
Q04251 arabidopsis
Q84hm9 streptomyces
Q94g13 oryza sativ
Q9e4m7 streptococ
Q91g25 oryza sativ
Q941f2 oryza sativ
Q84mw8 oryza sativ
Q8myj1 caenorhabdi
Q9bwx0 mus musculu
Q94772 caulobact
Q94j12 oryza sativ
Q7nd05 mus musculu
Q7b000 prochloroco
Q858a0 mycobacteri
Q9xe2a oryza sativ
Q9rzt5 oryza sativ
Q8t8f8 methanosarc
Q9bw69 mus musculu

ALIGNMENTS

[illegible]

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	91.5	79.6	287	13	Q93503	Q93503 xenopus lae	
2	91.5	79.6	287	13	Q8AVY4	Q8AVY4 xenopus lae	
3	72	62.6	232	13	Q7SXW9	Q7SXW9 brachydanio	
4	53	46.1	356	10	Q9AS10	Q9AS10 oryza sativ	
5	50	43.5	314	2	O68475	O68475 xanthomonas	
6	50	43.5	399	16	Q8U7M4	Q8U7M4 agrobacteri	
7	49	42.6	2785	17	Q8ZVB9	Q8ZVB9 pyrobaculum	
8	48	41.7	261	16	Q8ZH26	Q8ZH26 yersinia pe	
9	48	41.7	287	16	Q8FSA8	Q8FSA8 corynebacte	
10	48	41.7	334	16	Q89CP6	Q89CP6 bradyrhizob	
11	48	41.7	413	16	Q81Y95	Q81Y95 bacillus an	
12	48	41.7	465	10	Q9LG84	Q9LG84 oryza sativ	
13	48	41.7	527	10	Q7X711	Q7X711 oryza sativ	
14	48	41.7	790	2	Q9KW22	Q9KW22 xanthomonas	
15	48	41.7	1058	10	Q9LEZ5	Q9LEZ5 arabidopsis	
16	47.5	41.3	386	2	Q93J05	Q93J05 rhodococcus	

Query Match 79.6%; Score 91.5; DB 13; Length 287;

Matches	21;	Conservative	1;	Mismatches	2;	Indels	1;	Gaps	1;
----------------	------------	---------------------	-----------	-------------------	-----------	---------------	-----------	-------------	-----------

1 GAOFSKTAAKGEAA-AERPGEAAVA 24

db 2 GA0FSKTA0KGEAA0AEKPGEA0VPA 26


```

DR  PIR, AC3089; AC3089.
DR  PIR, G98197; G98197.
DR  InterPro, IPR000253; FHA.
DR  InterPro, IPR008984; SMAD_FHA.
DR  Pfam, PF00498; FHA; 1.
DR  SMART, SM00240; FHA; 1.
DR  PROSITE, PS00006; FHA_DOMAIN; 1.
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 399 AA; 43398 MW; 3719D2FFEC234E4 CRC64;

Query Match 43.5%; Score 50; DB 16; Length 399;
Best Local Similarity 50.0%; Pred. No. 57;
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 QFGKTAAKGEAAERPGEEAAVA 24
    ||:||||:|:|
DB 367 QFLQRAARGEAASARETDATAA 388

RESULT 7
Q8ZYB9
ID Q8ZYB9 PRELIMINARY; PRT; 2785 AA.
AC Q8ZYB9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PAREP2b.
DE DE
GN PAE0850.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
NCBI_TaxID=13773;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
RE EMBL; AE009790; AAL63076.1; -.
KW Complete proteome.
SQ SEQUENCE 2785 AA; 300989 MW; 1E144809E1D59C2D CRC64;

Query Match 42.6%; Score 49; DB 17; Length 2785;
Best Local Similarity 57.9%; Pred. No. 57e+02;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 SKTAAKGEAAERPGEEAAV 23
    ::|||||:|
DB 1343 AELAFKGYAARPPAEAL 1361

RESULT 8
Q8ZH26
ID Q8ZH26 PRELIMINARY; PRT; 261 AA.
AC Q8ZH26;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative flagellar basal-body rod protein (Flagellar biosynthesis
DE protein, cell-distal portion of basal-body rod).
GN FLGG OR YP00728 OR Y3450.
OS Versinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Versinia.
NCBI_TaxID=632;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,

```


RA Holtzapfel E.K., Oktad O.A., Helgason E., Rilstone J., Wu M.,
 RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,
 RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
 RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
 RA Benton J.L., Mahmoud Y., Jiang L., Hance I.R., Weidman J.F.,
 RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
 RA Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salzberg S.L.,
 RA Thomson B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,
 RA Fraser C.M.;
 RT "The genome sequence of *Bacillus anthracis* Ames and comparison to
 RT closely related bacteria";
 RL Nature 423:81-86(2003);
 DR EMBL; AE017035; AAP27411.1; -;
 DR TIGR; BA3660; -;
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR001478; PDZ.
 DR InterPro; IPR001254; Peptidase S1.
 DR InterPro; IPR008353; Peptidase S1B tx.
 DR InterPro; IPR001940; Peptidase_S1C.
 DR Pfam; PF00595; PDZ; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR01774; EXFOLTOXIN.
 DR PRINTS; PR00834; PROTEASES2C.
 DR SMART; SM00228; PDZ; 1.
 KW Protease; Complete proteome.
 SQ SEQUENCE 413 AA; 43899 MW; 77CCAL1656A8B4555 CRC64;

Query Match 41.7%; Score 48; DB 16; Length 413;
 Best Local Similarity 41.7%; Pred. No. 1.1e+02;
 Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GAQFSKTAAGKGAARPERGEAAVA 24

Db 181 GANYNKVATLGDSSKRAGEAAIA 204

RESULT 12

ID Q9LGS4 PRELIMINARY; PRT; 465 AA.
 AC Q9LGS4
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE P0710805.11 protein.
 GN P0710805.11.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoideae; Oryza.
 OC NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 RT clone:P0710E05";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP002743; BA999426.1; -;
 DR Gramene; Q9LGS4; -;
 DR InterPro; IPR000048; IQ_region.
 DR Pfam; PF00612; IQ; 2.
 DR SMART; SM00015; IQ; 2.
 SQ SEQUENCE 465 AA; 49651 MW; 82B7D93B2B4E55CE CRC64;

Query Match 41.7%; Score 48; DB 10; Length 465;
 Best Local Similarity 47.8%; Pred. No. 1.3e+02;
 Matches 11; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 2 AQFSKTAAGKGAARPERGEAAVA 24

Db 43 AKSRSDSTEGEAAAVGNAAIA 65
 ID Q7X711 PRELIMINARY; PRT; 527 AA.
 AC Q7X711
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE OSJNBAA0042115.3 protein (OJ000315_02.21 protein).
 GN OSJNBAA0042115.3 OR OJ000315_02.21.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoideae; Oryza.
 OC NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F.,
 RA Jia J., Yin H.F., Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y.,
 RA Shao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W.,
 RA Sheng H.H., Gu J.L., Chen S.T., Ni L., Zhu F.H., Han B., Feng Q.,
 RA Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X., Liu Y.L., Mu J., Yu Z.,
 RA Chen L., Fan D.L., Wang Q.J., Zhang L., Lu Y.Q., Yu S.L., Liu X.H.,
 RA Lu T.T., Zhang Y.J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,
 RA Qian Y.M., Ying K., Zhou B., Chen Z.H., Hao P., Zhang L., Wu M.,
 RA Zhang R.Q., Guan J.P., Hong G.F.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL731641; CAE04881.1; -;
 DR EMBL; AL731639; CAE05376.1; -;
 SQ SEQUENCE 527 AA; 56268 MW; 805B3E63409B269D CRC64;

Query Match 41.7%; Score 48; DB 10; Length 527;
 Best Local Similarity 55.0%; Pred. No. 1.5e+02;
 Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 5 SKTAAGKGAARPERGEAAVA 24

Db 139 SSTAAGSGGARRNGAAIA 158

RESULT 14

ID Q9KW22 PRELIMINARY; PRT; 790 AA.
 AC Q9KW22
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE HRPF.
 GN HRPF.
 OS Xanthomonas oryzae (pv. oryzae).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OC NCBI_TaxID=64187;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF 311018;
 RA MEDLINE=21303248; PubMed=11410350;
 RA Ochiai H., Inoue Y., Hasebe A., Kaku H.;
 RT "Construction and characterization of a Xanthomonas oryzae pv. oryzae
 RT bacterial artificial chromosome library."
 RL FEMS Microbiol. Lett. 200:59-65 (2001).
 DR EMBL; AB045312; BAB07869.1; -;
 DR InterPro; IPR008718; NOLX.
 DR Pfam; PF05819; NOLX; 1.
 SQ SEQUENCE 790 AA; 84883 MW; 73FD1F71106E56B8 CRC64;

Query Match 41.7%; Score 48; DB 2; Length 790;
 Best Local Similarity 59.1%; Pred. No. 2.2e+02;
 Matches 13; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

```
QY 5 SKTAAK--GEAAARPGEAAY 24
DB 689 AKGAAGAGKTAARPSAAFA 710

RESULT 15
Q9LEZ5
ID Q9LEZ5 PRELIMINARY; PRT; 1058 AA.
AC Q9LEZ5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN T22D6_50.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Terryn N., Ardiles W., Buysshaert C., Dasseville R.,
RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,
RA Villarroel R., Gielen J., Van Montagu M., Bancroft I., Mewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL357612; CAB93712.1; -.
DR FIR; T50496; T50496.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.
DR GO; GO:0016787; F:Hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR01410; DEAD.
DR InterPro; IPR01650; Helicase_C.
DR InterPro; IPR00626; Ubiquitin.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SMO0487; DEXDc; 1.
DR SMART; SMO0490; HELICc; 1.
DR PROSITE; PS0053; UBIQUITIN_2; 1.
KW Hypothetical protein; ATP-binding; Helicase; Hydrolase.
SQ SEQUENCE 1058 AA; 116856 MW; D95C267BF2672061 CRC64;

Query Match 41.7%; Score 48; DB 10; Length 1058;
Best Local Similarity 52.6%; Pred. No. 2.9e+02;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 5 SKTAAGGEAAARPGEAAY 23
DB 532 SKVSSSEAAADKPSGAAY 550

Search completed: October 7, 2004, 17:35:31
Job time : 38.5 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:32:02 ; Search time 50.5 seconds
(without alignments)
134.280 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GAQFSKTAAGKGEAAARPGEAAVA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	100.0	24	3	AAY95896 MANS pept
2	115	100.0	24	6	ABP97765 Myristoyl
3	115	100.0	165	7	ADe80715 Microsate
4	115	100.0	182	7	ADe80716 Microsate
5	115	100.0	308	7	ADe60419 Rat Prote
6	115	100.0	308	7	ADe60423 Rat Prote
7	115	100.0	308	7	ADe60493 Rat Prote
8	115	100.0	308	7	ADe45848 Rat Prote
9	115	100.0	330	2	AAR05528 High dens
10	115	100.0	332	3	AAY95898 Human myr
11	115	100.0	332	3	AAY95899 Human myr
12	115	100.0	332	7	ADe80714 Microsate
13	111	96.5	309	7	ADD47655 Rat Prote
14	49	42.6	380	4	ABU18252 Protein e
15	48	41.7	413	6	ABU18252 Protein e
16	47	40.9	104	7	ADe10634 Structura
17	47	40.9	341	4	ABB60616 Drosophil
18	47	40.9	410	6	ABU18611 Protein e
19	47	40.9	559	6	ABU41740 Protein e
20	46.5	40.4	1128	3	AAG29155 Arabidops
21	46.5	40.4	1206	3	AAG29154 Arabidops
22	46.5	40.4	1293	3	AAG29153 Arabidops
23	46	40.0	916	4	ABB71090 Drosophil
24	45	39.1	89	4	AAM21657 Peptide #
25	45	39.1	89	4	ABB44016 Peptide #

ALIGNMENTS

RESULT 1

AAY95896

ID AAY95896 standard; peptide; 24 AA.

XX AC AAY95896;

XX AC AAY95896;

DT 20-NOV-2000 (first entry)

DE MANS peptide inhibitor of MARCKS-related mucus secretion.

XX MANS; MARCKS; myristoylated alanine-rich C kinase substrate; human;

KW mucus secretion; inhibitor; bronchitis; asthma; cystic fibrosis;

KW chronic obstructive pulmonary disease; pneumonia; empysema; influenza;

KW rhinitis; therapy.

XX Homo sapiens.

XX WO200050062-A2.

PD 31-AUG-2000.

PF 24-FEB-2000; 2000WO-US005050.

XX 24-FEB-1999; 99US-00256154.

XX (UYN-) UNIV NORTH CAROLINA STATE.

PI Li Y, Martin LD, Adler KB;

XX WPI; 2000-572036/53.

Regulating mucus secretion by a mucus-secreting cell, useful for treating e.g. bronchitis, asthma or pneumonia, by administering a compound that inhibits or enhances myristoylated alanine-rich C-kinase substrate protein.
Claim 12; Page 40; 66pp; English.

The present sequence is that of MANS peptide, comprising the N-terminal region of human myristoylated alanine-rich C kinase substrate MARCKS protein (see AAY95898), a major cellular substrate for protein kinase S. MANS peptide inhibits secretion of mucus from mucus membranes and mucous-secreting cells, including human airway epithelial cells. It is suggested to block attachment of MARCKS protein to the mucin granule, thus blocking or inhibiting the release of mucin granules and the secretion of mucus by the cell. The invention relates to methods and compounds for decreasing mucus secretion, particularly in the airways. Such compounds include MANS peptide and antisense oligonucleotides to MARCKS. They are useful in

CC inhibiting mucus secretion in conditions such as bronchitis, cystic
CC fibrosis, chronic obstructive pulmonary disease, asthma, emphysema,
CC pneumonia, influenza, rhinitis and the common cold
XX Sequence 24 AA;
SQ

Query Match 100.0%; Score 115; DB 3; Length 24;
Best Local Similarity 100.0%; Pred. No. 9.6e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAARKGEAAERPGGEAAVA 24
DB 1 GAQFSKTAARKGEAAERPGGEAAVA 24

RESULT 2
ID ABP97765 standard; peptide; 24 AA.
XX
AC ABP97765;
XX
DT 11-AUG-2003 (first entry)
XX
DE Myristoylated N-terminal sequence (MANS) peptide.
XX
KW Myristoylated N-terminal sequence; MANS; MARCKS; mucus secretion;
KW inflammatory mediator; inflammation; respiratory disease; asthma;
KW chronic bronchitis; chronic obstructive pulmonary disease; COPD;
KW bowel disease; irritable bowel syndrome; Crohn's disease;
KW ulcerative colitis; skin disease; rosacea; eczema; psoriasis; acne;
KW autoimmune disease; pain; arthritis; cystic fibrosis.
XX
OS Synthetic.

Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal myristate chain attached"
FT
XX
XX WO2003000027-A2.
XX
XX 03-JAN-2003.
XX
XX 26-JUN-2002; 2002WO-US022270.
XX
XX 26-JUN-2001; 2001US-0300933P.
XX
XX (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX Martin LD, Adler KB, Li Y;
XX
XX WPI; 2003-278239/27.
XX
XX Method of regulating inflammation comprises administering a composition
XX comprising a MANS peptide or an active fragment thereof.
XX
XX Disclosure; Page 26; 54pp; English.
XX
XX The present sequence represents a myristoylated N-terminal sequence
XX (MANS) peptide. The MANS peptide is identical to the first 24 amino acids
XX of MARCKS, and mediates insertion of MARCKS into membranes. The MANS
XX peptide inhibits both mucus secretion and inflammatory mediators. The
XX specification describes a method of regulating inflammation. The method
XX comprises administering a composition comprising a MANS peptide. The
XX method is useful for treating inflammation caused by respiratory diseases
XX (e.g. asthma, chronic bronchitis and chronic obstructive pulmonary
XX disease (COPD), bowel diseases (e.g. irritable bowel syndrome, Crohn's
XX disease and ulcerative colitis), skin diseases (e.g. rosacea, eczema,
XX psoriasis and severe acne), autoimmune diseases and pain syndromes,
XX arthritis and cystic fibrosis
XX
XX Sequence 24 AA;
SQ

Query Match 100.0%; Score 115; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 9.6e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 9.6e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAARKGEAAERPGGEAAVA 24
DB 1 GAQFSKTAARKGEAAERPGGEAAVA 24

RESULT 3
ADE80715
ID ADE80715 standard; protein; 165 AA.
XX
AC ADE80715;
XX
DT 29-JAN-2004 (first entry)
XX
DE Microsatellite related MACS (-1) ORF amino acid sequence.
XX
KW frameshift mutation; microsatellite; cytostatic; neuroprotective;
KW vasotrophic; vaccine; gene therapy; neurodegenerative disorder;
KW vascular disease; cancer.
XX
OS Unidentified.
XX
FN WO2003087162-A2.
XX
PD 23-OCT-2003.
XX
XX 17-APR-2003; 2003WO-EP004083.
XX
XX 18-APR-2002; 2002EP-00008771.
XX
XX 18-APR-2002; 2002EP-00008773.
XX
XX 18-APR-2002; 2002EP-00008774.
XX
XX (MTMM-) MTM LAB AG.
XX
XX Von Knebel Doebritz M, Gebert J, Linnebacher M, Woerner S;
XX Ridder R, Bork P, Yuan YP;
XX
XX WPI; 2003-845308/78.
XX
XX New nucleic acid, useful in preparing a composition for diagnosing or
XX treating disorders associated with frameshift mutations in coding
XX microsatellite regions, e.g., neurodegenerative disorder, vascular
XX disease or cancer.
XX
XX Claim 3; Fig 2; 62pp; English.
XX
XX The present invention describes a nucleic acid sequence (I) which encodes
XX a polypeptide consisting of TAF1B, MACS, UVRAG, ELAVL3, TCF6L1, ABCF1,
XX AIM2, CHD2, FL J11053, KIAA1052, ACVR2 or HT001 having a frameshift
XX mutation. Also described: (1) a frameshift polypeptide (II); (2) a method
XX for treating disorders associated with frameshift mutations in coding
XX microsatellites; (3) a pharmaceutical composition comprising the nucleic
XX acid and/or polypeptide; (4) a method for detecting a disorder associated
XX with frameshift mutations in coding microsatellite regions; (5) a
XX diagnostic or research kit for detecting a disorder associated with
XX frameshift mutations in coding microsatellite regions, comprising the
XX nucleic acid and/or frameshift polypeptide; and (6) a method for treating
XX disorders associated with peptides arising from frameshift mutations in
XX coding microsatellite regions in individuals. (I) and (II) have
XX cytosolic, neuroprotective and vasotropic activities, and can be used in
XX vaccines and in gene therapy. The nucleic acid (I) or frameshift
XX polypeptide (II) can be used in detecting disorders associated with
XX frameshift mutations in coding microsatellite regions or in preparing
XX pharmaceutical compositions for treating disorders associated with
XX frameshift mutations in coding microsatellite regions, e.g.,
XX neurodegenerative disorder, vascular disease, cancer or precursory stages
XX of cancer. The present sequence represents a protein used in the
XX exemplification of the present invention.
XX
XX Sequence 165 AA;
SQ

Query Match 100.0%; Score 115; DB 7; Length 165;
 Best Local Similarity 100.0%; Pred. No. 7.5e-10;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAARPGGAAVA 24
 |||||
 Db 2 GAQFSKTAAGGAAARPGGAAVA 25

RESULT 4

ADBE0716
 ID ADBE0716 standard; protein; 182 AA.

XX AC ADBE0716;

XX DT 29-JAN-2004 (first entry)

XX DE Microsatellite related MACS (+1)/(-2) ORF amino acid sequence.

XX DE frameshift mutation; microsatellite; cytostatic; neuroprotective;
 KW vasotopic; vaccine; gene therapy; neurodegenerative disorder;
 KW vascular disease; cancer.

XX OS Unidentified.

XX SN WO2003087162-A2.

XX FN 23-OCT-2003.

XX PD 17-APR-2003; 2003WO-EP004083.

XX PF 18-APR-2002; 2002EP-00008771.

XX PR 18-APR-2002; 2002EP-00008773.

XX PR 18-APR-2002; 2002EP-00008774.

XX PA (MTMM-) MTM LAB AG.

XX PI Von Knebel Doberitz M, Gebert J, Linnebacher M, Woerner S;
 PI Ridder R, Bork P, Yuan YP;

XX DR WPI; 2003-845308/78.

XX PT New nucleic acid, useful in preparing a composition for diagnosing or
 PT treating disorders associated with frameshift mutations in coding
 PT microsatellite regions, e.g., neurodegenerative disorder, vascular
 PT disease or cancer.

XX PS Claim 3; Fig 2; 62pp; English.

CC The present invention describes a nucleic acid sequence (I) which encodes
 CC a polypeptide consisting of TAF1B, MACS, UVRAG, ELAVL3, TCF6L1, ABCF1,
 CC AIM2, CHD2, FL J11053, KIAA1052, ACVR2 or HT001 having a frameshift
 CC mutation. Also described: (1) a frameshift polypeptide (II); (2) a method
 CC for treating disorders associated with frameshift mutations in coding
 CC microsatellites; (3) a pharmaceutical composition comprising the nucleic
 CC acid and/or polypeptide; (4) a method for detecting a disorder associated
 CC with frameshift mutations in coding microsatellite regions; (5) a
 CC diagnostic or research kit for detecting a disorder associated with
 CC frameshift mutations in coding microsatellite regions, comprising the
 CC nucleic acid and/or frameshift polypeptide; and (6) a method for treating
 CC disorders associated with peptides arising from frameshift mutations in
 CC coding microsatellite regions in individuals. (I) and (II) have
 CC cytostatic, neuroprotective and vasotropic activities, and can be used in
 CC vaccines and in gene therapy. The nucleic acid (I) or frameshift
 CC polypeptide (II) can be used in detecting disorders associated with
 CC frameshift mutations in coding microsatellite regions or in preparing
 CC pharmaceutical compositions for treating disorders associated with
 CC frameshift mutations in coding microsatellite regions, e.g.,
 CC neurodegenerative disorder, vascular disease, cancer or precursory stages
 CC of cancer. The present sequence represents a protein used in the
 CC exemplification of the present invention.

XX SQ Sequence 182 AA;

Query Match 100.0%; Score 115; DB 7; Length 182;
 Best Local Similarity 100.0%; Pred. No. 8.3e-10;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAARPGGAAVA 24
 |||||
 Db 2 GAQFSKTAAGGAAARPGGAAVA 25

RESULT 5

ADBE0419

ID ADBE0419 standard; protein; 308 AA.

XX AC ADBE0419;

XX DT 29-JAN-2004 (first entry)

XX DE Rat Protein P30009, SEQ ID NO 6328.

XX DE Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX SN WO2003016475-A2.

XX FN 27-FEB-2003.

XX PD 14-AUG-2002; 2002WO-US025765.

XX PF 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GEHO) GEN HOSPITAL CORP.

XX FA (FARB) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX DR WPI; 2003-268312/26.

XX PR GENBANK; P30009.

XX PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNI)), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 308 AA;

SQ Query Match 100.0%; Score 115; DB 7; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAQFSKTAAGGAAERPGGAAVA 24
Db 1 GAQFSKTAAGGAAERPGGAAVA 24

RESULT 6
ADE60423
ID ADE60423 standard; protein; 308 AA.

XX AC ADE60423;

XX DT 29-JAN-2004 (first entry)

XX DE Rat Protein P30009, SEQ ID NO 6332.

XX XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX XX WO2003016475-A2.

XX FN 27-FEB-2003.

XX XX 14-AUG-2002; 2002WO-US025765.

XX PR 01-NOV-2001; 2001US-0312147P.

XX PR 26-NOV-2001; 2001US-0346382P.

XX XX 26-NOV-2001; 2001US-0333347P.

XX XX (GEHO) GEN HOSPITAL CORP.

XX PA (FARB) BAYER AG.

XX XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX DR GENBANK; P30009.

XX PT New composition comprising two or more isolated polypeptides, useful for

XX PS preparing a medicament for treating pain in an animal.

XX XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 308 AA;

Query Match 100.0%; Score 115; DB 7; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAQFSKTAAGGAAERPGGAAVA 24
Db 1 GAQFSKTAAGGAAERPGGAAVA 24

RESULT 7
ADE60493
ID ADE60493 standard; protein; 308 AA.

XX AC ADE60493;

XX DT 29-JAN-2004 (first entry)

XX DE Rat Protein P30009, SEQ ID NO 6402.

XX XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX XX WO2003016475-A2.

XX PD 27-FEB-2003.

XX XX 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX XX (GEHO) GEN HOSPITAL CORP.

XX PA (FARB) BAYER AG.

XX XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX DR GENBANK; P30009.

XX PT New composition comprising two or more isolated polypeptides, useful for

XX PS preparing a medicament for treating pain in an animal.

XX XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating

CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 308 AA;

Query Match 100.0%; Score 115; DB 7; Length 308;
 Best Local Similarity 100.0%; Pred. No. 1.5e-09;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAERPGGAAVA 24
 |||||
 Db 1 GAQFSKTAAGGAAERPGGAAVA 24

RESULT 8

ADD45848
 ID ADD45848 standard; protein; 308 AA.

AC ADD45848;

AC 29-JAN-2004 (first entry)

DE Rat Protein P30009, SEQ ID NO 11518.

Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX (GSHO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

DR GENEANK; P30009.

XX New composition comprising two or more isolated polypeptides, useful for
 preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the

CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 308 AA;

Query Match 100.0%; Score 115; DB 7; Length 308;
 Best Local Similarity 100.0%; Pred. No. 1.5e-09;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAERPGGAAVA 24
 |||||
 Db 1 GAQFSKTAAGGAAERPGGAAVA 24

RESULT 9

AAR05528

ID AAR05528 standard; protein; 330 AA.

XX AAR05528;

XX 25-MAR-2003 (revised)

DT 23-OCT-1990 (first entry)

XX High density lipoprotein (HDL) binding protein.

XX High density lipoprotein; HDL-binding protein; atherosclerosis;
 KW hypercholesterolaemia; ds.

XX Homo sapiens.

XX WO9005744-A.

XX 31-MAY-1990.

XX 18-NOV-1988; 88US-00273388.

XX 18-NOV-1988; 88US-00273388.

XX 05-MAY-1989; 89US-00347855.

XX (UNIW) UNIV WASHINGTON.

XX (ZYMO) ZYMOGENETICS INC.

XX Oram JE, McKnight GL, Hart CE, Curtis DA;

XX WPI; 1990-193405/25.

XX N-PSDB; AAQ04784.

XX New mammalian proteins binding high density lipoprotein sub-class 3 - DNA
 encoding them and derived antibodies, for screening potentially
 therapeutic HDL analogues and for diagnosing risk of atherosclerosis.

PS Claim 4; Fig 1A-D; 79pp; English.

XX The protein product may be used to raise Abs, and the cDNA to create
 CC probes, both useful in screening for HDL analogues, agonists and
 CC antagonists, and in identifying abnormalities in the HDL binding/receptor
 CC pathway. HDL analogues can be used in treating hypercholesterolaemia and
 CC atherosclerosis (Updated on 25-MAR-2003 to correct PR field.) (Updated on
 CC 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI
 CC field.)

```
SQ Sequence 330 AA;
Query Match 100.0%; Score 115; DB 2; Length 330;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAQFSKTAARKGEAAERPGGEAAVA 24
Db 2 GAQFSKTAARKGEAAERPGGEAAVA 25

RESULT 10
AAY95898
ID AAY95898 standard; protein; 332 AA.
AC AAY95898;
XX 20-NOV-2000 (first entry)
XX Human myristoylated alanine-rich C kinase substrate MARCKS.
XX MARCKS; myristoylated alanine-rich C kinase substrate; human;
XX mucus secretion; inhibitor; bronchitis; asthma; cystic fibrosis;
XX chronic obstructive pulmonary disease; pneumonia; emphysema; influenza;
XX rhinitis; therapy.
XX Homo sapiens.
XX Key Location/Qualifiers
FH Peptide 2..25
FT /note= "MANS peptide of AAY95898"
FT Misc-difference 84
FT /note= "Ser in sequence of AAY95899"
FT Misc-difference 119
FT /note= "Ala in sequence of AAY95899"
FT Peptide 152..176
FT /note= "MA-PSD peptide of AAY95897"
XX WO200050062-A2.
XX 31-AUG-2000.
XX 24-FEB-2000; 2000WO-US005050.
XX 24-FEB-1999; 99US-00256154.
XX (UYN-) UNIV NORTH CAROLINA STATE.
XX Li Y, Martin LD, Adler KB;
XX WPI; 2000-572036/53.
XX N-PSDB; AAA50339.
XX Regulating mucus secretion by a mucus-secreting cell, useful for treating
XX e.g. bronchitis, asthma or pneumonia, by administering a compound that
XX inhibits or enhances myristoylated alanine-rich C-kinase substrate
XX protein.
XX Claim 3; Page 42-43; 66pp; English.
XX The present sequence is that of human myristoylated alanine-rich C kinase
XX substrate MARCKS protein, a major cellular substrate. The invention
XX relates to methods of inhibiting mucus secretion by a mucus-secreting
XX cell by administering a compound that inhibits MARCKS protein-related
XX mucus secretion. Such compounds include active fragments of MARCKS
XX protein such as MANS peptide (see AAY95897) and MA-PSD peptide (see
XX AAY95897), which corresponds to a phosphorylation site of MARCKS. The
XX inhibitor compounds can be used to treat conditions such as bronchitis,
XX cystic fibrosis, chronic obstructive pulmonary disease, asthma,
XX emphysema, pneumonia, influenza, rhinitis and the common cold. An
XX alternative sequence for MARCKS is provided in AAY95899, which differs
XX from the present sequence at 2 amino acid residues, Ala-84 (Ser) and Pro-
XX 119 (Ala)
```

```
XX SQ Sequence 332 AA;
Query Match 100.0%; Score 115; DB 3; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAQFSKTAARKGEAAERPGGEAAVA 24
Db 2 GAQFSKTAARKGEAAERPGGEAAVA 25

RESULT 11
AAY95899
ID AAY95899 standard; protein; 332 AA.
XX AAY95899;
XX 20-NOV-2000 (first entry)
XX Human myristoylated alanine-rich C kinase substrate MARCKS.
XX MARCKS; myristoylated alanine-rich C kinase substrate; human;
XX mucus secretion; inhibitor; bronchitis; asthma; cystic fibrosis;
XX chronic obstructive pulmonary disease; pneumonia; emphysema; influenza;
XX rhinitis; therapy.
XX Homo sapiens.
XX Key Location/Qualifiers
FH Peptide 2..25
FT /note= "MANS peptide of AAY95896"
FT Misc-difference 84
FT /note= "Ala in sequence of AAY95898"
FT Misc-difference 119
FT /note= "Pro in sequence of AAY95898"
FT Peptide 152..176
FT /note= "MA-PSD peptide of AAY95897"
XX WO200050062-A2.
XX 31-AUG-2000.
XX 24-FEB-2000; 2000WO-US005050.
XX 24-FEB-1999; 99US-00256154.
XX (UYN-) UNIV NORTH CAROLINA STATE.
XX Li Y, Martin LD, Adler KB;
XX WPI; 2000-572036/53.
XX N-PSDB; AAA50340.
XX Regulating mucus secretion by a mucus-secreting cell, useful for treating
XX e.g. bronchitis, asthma or pneumonia, by administering a compound that
XX inhibits or enhances myristoylated alanine-rich C-kinase substrate
XX protein.
XX Disclosure; Page 46-47; 66pp; English.
XX The present sequence is that of human myristoylated alanine-rich C kinase
XX substrate MARCKS protein, a major cellular substrate. The invention
XX relates to methods of inhibiting mucus secretion by a mucus-secreting
XX cell by administering a compound that inhibits MARCKS protein-related
XX mucus secretion. Such compounds include active fragments of MARCKS
XX protein such as MANS peptide (see AAY95897) and MA-PSD peptide (see
XX AAY95897), which corresponds to a phosphorylation site of MARCKS. The
XX inhibitor compounds can be used to treat conditions such as bronchitis,
XX cystic fibrosis, chronic obstructive pulmonary disease, asthma,
XX emphysema, pneumonia, influenza, rhinitis and the common cold. An
XX alternative sequence for MARCKS is provided in AAY95898, which differs
XX from the present sequence at 2 amino acid residues, Ser-84 (Ala) and Ala-
```


CC 119 (Pro)
XX Sequence 332 AA;
SQ Query Match 100.0%; Score 115; DB 3; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAARPGGAAVA 24
|||
Db 2 GAQFSKTAAGGAAARPGGAAVA 25
|||

RESULT 12
ADE0714
ID ADE0714 standard; protein; 332 AA.
XX AC ADE0714;
XX 29-JAN-2004 (first entry)
XX Microsatellite related MACS wt ORF amino acid sequence.
XX frameshift mutation; microsatellite; cytostatic; neuroprotective;
KW vasotropic; vaccine; gene therapy; neurodegenerative disorder;
KW vascular disease; cancer.
XX Unidentified.
XX WO2003087162-A2.
XX 23-OCT-2003.
XX 17-APR-2003; 2003WO-EP004083.
XX 18-APR-2002; 2002EP-00008771.
PR 18-APR-2002; 2002EP-00008773.
PR 18-APR-2002; 2002EP-00008774.
XX (MTWM-) MTM LAB AG.
XX Von Knebel Doeberitz M, Gebert J, Linnebacher M, Woerner S;
PI Ridder R, Bork P, Yuan YP;
XX WPI; 2003-845308/78.
XX New nucleic acid, useful in preparing a composition for diagnosing or
PT treating disorders associated with frameshift mutations in coding
PT microsatellite regions, e.g., neurodegenerative disorder, vascular
PT disease or cancer.
XX Claim 3; Fig 2; 62pp; English.

CC The present invention describes a nucleic acid sequence (I) which encodes
CC a polypeptide consisting of TAF1B, MACS, UVRAG, ELAVL3, TCF6L1, ABCF1,
CC AIM2, CHD2, FLJ11053, KIAA1052, ACVR2 or HT001 having a frameshift
CC mutation. Also described: (1) a frameshift polypeptide (II); (2) a method
CC for treating disorders associated with frameshift mutations in coding
CC microsatellites; (3) a pharmaceutical composition comprising the nucleic
CC acid and/or polypeptide; (4) a method for detecting a disorder associated
CC with frameshift mutations in coding microsatellite regions; (5) a
CC diagnostic or research kit for detecting a disorder associated with
CC frameshift mutations in coding microsatellite regions, comprising the
CC nucleic acid and/or frameshift polypeptide; and (6) a method for treating
CC disorders associated with peptides arising from frameshift mutations in
CC coding microsatellite regions in individuals. (I) and (II) have
CC cytostatic, neuroprotective and vasotropic activities, and can be used in
CC vaccines and in gene therapy. The nucleic acid (I) or frameshift
CC polypeptide (II) can be used in detecting disorders associated with
CC frameshift mutations in coding microsatellite regions or in preparing
CC pharmaceutical compositions for treating disorders associated with
CC frameshift mutations in coding microsatellite regions, e.g.,
CC neurodegenerative disorder, vascular disease, cancer or precursory stages

CC of cancer. The present sequence represents a protein used in the
CC exemplification of the present invention.
XX
SQ Sequence 332 AA;
Query Match 100.0%; Score 115; DB 7; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAARPGGAAVA 24
|||
Db 2 GAQFSKTAAGGAAARPGGAAVA 25
|||

RESULT 13
ADD47655
ID ADD47655 standard; protein; 309 AA.
XX AC ADD47655;
XX 29-JAN-2004 (first entry)
XX Rat Protein NP_032564, SEQ ID NO 13351.
DE Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX Rattus norvegicus.
XX WO2003016475-A2.
XX 27-FEB-2003.
XX 14-AUG-2002; 2002WO-US025765.
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX Woolf C, D'urso D, Befort K, Costigan M;
PI WPI; 2003-268312/26.
DR GENBANK; NP_032564.
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
PT Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 309 AA;
 SQ

Query Match 96.5%; Score 111; DB 7; Length 309;
 Best Local Similarity 95.8%; Pred. No. 6.1e-09;
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAQFSKTAAGGAAARPGRAAVA 24
 |||||
 Db 2 GAQFSKTAAGGAAARPGRAAVA 25

RESULT 14
 ABG07150
 ID ABG07150 standard; protein; 380 AA.

XX AC ABG07150;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #7141.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSEQ-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS71337.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 20; SEQ ID NO 37509; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have application in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 380 AA;

Query Match 42.6%; Score 49; DB 4; Length 380;
 Best Local Similarity 68.8%; Pred. No. 35;
 Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 5 SKTAAKGEEAAERPGGE 20
 |||||
 Db 333 SKTAAKGEEAAERPGSK 348

RESULT 15
 ABU18252
 ID ABU18252 standard; protein; 413 AA.

XX AC ABU18252;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #3779.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Bacillus anthracis.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR N-PSDB; ACA22122.

XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 46176; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway; (8)
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 413 AA;

Query Match 41.7%; Score 48; DB 6; Length 413;
 Best Local Similarity 41.7%; Pred. No. 55;
 Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAERPGGAAVA 24
 || : || : : : || : :
 Db 181 GANVKNVATLGDSSKIRAGEKAIA 204

Search completed: October 7, 2004, 17:33:56
 Job time : 56.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:35:42 ; Search time 84.5 Seconds
(without alignments)
91.399 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GAQFSKTAAGKGEAAERPGEAAVA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

```

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	100.0	24	US-10-180-753-1	Sequence 1, Appli
2	115	100.0	24	US-10-802-644-1	Sequence 1, Appli
3	115	100.0	389	US-10-131-410-172	Sequence 172, App
4	53	46.1	356	US-10-437-963-185520	Sequence 185520,
5	51	44.3	218	US-10-767-701-46204	Sequence 46204, A
6	51	44.3	1152	US-10-437-963-131104	Sequence 131104,
7	49.5	43.0	254	US-10-437-963-204840	Sequence 204840,
8	48.5	42.2	549	US-10-369-493-10346	Sequence 10346, A
9	48	41.7	148	US-10-767-701-37911	Sequence 37911, A
10	48	41.7	433	US-10-282-122A-46176	Sequence 46176, A
11	48	41.7	465	US-10-437-963-171398	Sequence 171398,
12	48	41.7	981	US-10-437-963-127070	Sequence 127070,
13	47.5	41.3	467	US-10-424-599-169015	Sequence 169015,
14	47	40.9	104	US-10-393-449-41	Sequence 41, Appl
15	47	40.9	104	US-10-177-725-41	Sequence 41, Appl

```

16 47 40.9 176 16 US-10-767-701-56335 Sequence 56335, A
17 47 40.9 200 16 US-10-767-701-35982 Sequence 35982, A
18 47 40.9 410 12 US-10-282-122A-46535 Sequence 46535, A
19 47 40.9 559 12 US-10-282-122A-69664 Sequence 69664, A
20 47 40.9 624 16 US-10-437-963-149882 Sequence 149882,
21 47 40.9 1350 16 US-10-647-196-27 Sequence 27, Appl
22 46.5 40.4 66 16 US-10-437-963-113367 Sequence 113367,
23 46 40.0 60 16 US-10-437-963-172410 Sequence 172410,
24 46 40.0 110 16 US-10-437-963-169526 Sequence 169526,
25 46 40.0 124 16 US-10-437-963-136066 Sequence 136066,
26 46 40.0 125 12 US-10-424-599-227273 Sequence 227273,
27 46 40.0 126 16 US-10-437-963-161220 Sequence 161220,
28 46 40.0 191 16 US-10-437-963-105384 Sequence 105384,
29 46 40.0 217 12 US-10-425-114-49555 Sequence 49555, A
30 46 40.0 324 16 US-10-437-963-144120 Sequence 144120,
31 46 40.0 439 16 US-10-437-963-188362 Sequence 188362,
32 46 40.0 612 16 US-10-437-963-185375 Sequence 185375,
33 45 39.1 89 9 US-09-864-761-42220 Sequence 42220, A
34 45 39.1 148 14 US-10-156-761-8289 Sequence 8289, Ap
35 45 39.1 239 16 US-10-437-963-191253 Sequence 191253,
36 45 39.1 255 16 US-10-437-963-204931 Sequence 204931,
37 45 39.1 469 16 US-10-437-963-126692 Sequence 126692,
38 45 39.1 1004 16 US-10-437-963-109426 Sequence 109426,
39 44.5 38.7 132 16 US-10-437-963-125382 Sequence 125382,
40 44.5 38.7 413 14 US-10-156-761-12219 Sequence 12219, A
41 44.5 38.7 440 15 US-10-369-493-8931 Sequence 8931, Ap
42 44 38.3 98 9 US-09-216-393-98 Sequence 98, Appl
43 44 38.3 98 14 US-10-321-856-98 Sequence 98, Appl
44 44 38.3 99 12 US-10-425-114-37338 Sequence 37338, A
45 44 38.3 111 16 US-10-767-701-53396 Sequence 53396, A

```

ALIGNMENTS

```

RESULT 1
US-10-180-753-1
; Sequence 1, Application US/10180753
; Publication No. US20030013652A1
; GENERAL INFORMATION:
; APPLICANT: Martin, Linda
; APPLICANT: Adler, Kenneth
; APPLICANT: Li, Yuehua
; TITLE OF INVENTION: BLOCKING PEPTIDE FOR INFLAMMATORY CELL SECRETION
; FILE REFERENCE: 5051.574
; CURRENT APPLICATION NUMBER: US/10/180.753
; CURRENT FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: N-terminal myristoylation
US-10-180-753-1

```

```

Query Match 100.0%; Score 115; DB 12; Length 24;
Best Local Similarity 100.0%; Pred. No. 8.6e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 GAQFSKTAAGKGEAAERPGEAAVA 24

Db 1 GAQFSKTAAGKGEAAERPGEAAVA 24

RESULT 2

US-10-802-644-1

; Sequence 1, Application US/10802644

```
/ Publication No. US20040180836A1
; GENERAL INFORMATION:
; APPLICANT: Martin, Linda
; APPLICANT: Adler, Kenneth
; APPLICANT: Li, Yuehua
; TITLE OF INVENTION: BLOCKING PEPTIDE FOR INFLAMMATORY CELL SECRETION
; FILE REFERENCE: 5051,574
; CURRENT APPLICATION NUMBER: US/10/802,644
; CURRENT FILING DATE: 2004-03-17
; PRIOR APPLICATION NUMBER: US/10/180,753
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: N-terminal myristoylation
US-10-802-644-1

Query Match          100.0%; Score 115; DB 16; Length 24;
Best Local Similarity 100.0%; Pred. No. 8.6e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAARPGAAVA 24
   |||||
DB 1 GAQFSKTAAGGAAARPGAAVA 24
   |||||

RESULT 3
US-10-131-410-172
; Sequence 172, Application US/10131410
; Publication No. US20030235915A1
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM TISSUE OF BREAST
; FILE REFERENCE: SCH-1763
; CURRENT APPLICATION NUMBER: US/10/131,410
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 09/646,673
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: PCT/DE99/00908
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 172
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-410-172

Query Match          100.0%; Score 115; DB 15; Length 389;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAARPGAAVA 24
   |||||
DB 59 GAQFSKTAAGGAAARPGAAVA 82
   |||||

RESULT 4
US-10-437-963-185520
```

```
/ Sequence 185520, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 185520
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_82408C.1.pep
US-10-437-963-185520

Query Match          46.1%; Score 53; DB 16; Length 356;
Best Local Similarity 75.0%; Pred. No. 17;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 TAAKGEAAARPGAA 22
   |||
DB 6 TAAKGEAAARPGAA 21
   |||
```

```
RESULT 5
US-10-767-701-46204
; Sequence 46204, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 46204
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C572_1.pep
US-10-767-701-46204

Query Match          44.3%; Score 51; DB 16; Length 218;
Best Local Similarity 55.6%; Pred. No. 19;
Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 6 KTAAGGAAARPGAAV 23
   : : : : : : : : : :
DB 33 RASAEAGAAAPGRRV 50
   : : : : : : : : : :

RESULT 6
US-10-437-963-131104
; Sequence 131104, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
```

APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 131104
LENGTH: 1152
TYPE: PRT
ORGANISM: Oryza sativa
NAME/KEY: unsure
LOCATION: (1)..(1152)
FEATURE:
OTHER INFORMATION: unsure at all Xaa locations
US-10-437-963-131104

Query Match 44.3%; Score 51; DB 16; Length 1152;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 11; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 QFSKTAAGGAAARPGGAAVA 24
DB 177 QFARTASSPASAPATAVA 198

RESULT 7

US-10-437-963-204840
Sequence 204840, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 204840
LENGTH: 254
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(254)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_99888C.1.pap
US-10-437-963-204840

Query Match 43.0%; Score 49.5; DB 16; Length 254;
Best Local Similarity 34.9%; Pred. No. 37;
Matches 15; Conservative 3; Mismatches 6; Indels 19; Gaps 1;

QY 1 GAQFSKTAAGGAAARP-----GEAAVA 24
DB 135 GSEASDGGEGGAAARPAWRRGRRGASGGGGGGEAAVA 177

RESULT 8

US-10-369-493-10346
Sequence 10346, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 10346
LENGTH: 549
TYPE: PRT
ORGANISM: Cytophaga hutchinsonii
US-10-369-493-10346

Query Match 42.2%; Score 48.5; DB 15; Length 549;
Best Local Similarity 54.2%; Pred. No. 1.2e+02;
Matches 13; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

QY 2 AQFSKTAAGGAA---AARPGGAA 22
DB 93 AEAPKTEAKAEVFKAEKPAEAA 116

RESULT 9

US-10-767-701-37911
Sequence 37911, Application US/10767701
Publication No. US20040172684A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 37911
LENGTH: 148
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(148)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-C40518_1.pap
US-10-767-701-37911

Query Match 41.7%; Score 48; DB 16; Length 148;
Best Local Similarity 73.3%; Pred. No. 34;
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 8 AAKGGAARPGGAA 22
DB 18 AAAAAAARPGGAA 32

RESULT 10

US-10-282-122A-46176
Sequence 46176, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu

```

RESULT 13
US-10-424-599-169015
; Sequence 169015, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 169015
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT MRT3847 123635C.1.pep

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:32:03 ; Search time 15 seconds
(without alignments)
82.602 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GAQFSKTAAGGAAERPGEAAVA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/2/iaa/5A COMB.pcp:*
- 2: /cgn2_6/prodata/2/iaa/5B COMB.pcp:*
- 3: /cgn2_6/prodata/2/iaa/6A COMB.pcp:*
- 4: /cgn2_6/prodata/2/iaa/6B COMB.pcp:*
- 5: /cgn2_6/prodata/2/iaa/6C COMB.pcp:*
- 6: /cgn2_6/prodata/2/iaa/6D COMB.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	100.0	309	2	US-08-405-175A-8
2	115	100.0	332	2	US-08-405-175A-5
3	111	96.5	309	2	US-08-405-175A-7
4	111	96.5	335	2	US-08-405-175A-6
5	104	90.4	281	2	US-08-405-175A-9
6	51	44.3	365	4	US-09-252-991A-32327
7	51	44.3	621	4	US-09-252-991A-19125
8	49	42.6	314	4	US-09-252-991A-26728
9	48	41.7	408	4	US-09-252-991A-31571
10	47	40.9	996	4	US-09-252-991A-28596
11	47	40.9	1257	4	US-09-252-991A-17290
12	46.5	40.4	416	4	US-09-252-991A-20341
13	46	40.0	207	4	US-09-252-991A-28823
14	46	40.0	380	4	US-09-252-991A-17603
15	46	40.0	758	4	US-09-252-991A-29974
16	45	39.1	172	4	US-09-252-991A-25750
17	45	39.1	694	4	US-09-252-991A-22279
18	44	38.3	98	4	US-09-216-393B-98
19	44	38.3	123	4	US-09-489-039A-9970
20	44	38.3	287	1	US-08-146-421-5
21	44	38.3	422	4	US-09-489-039A-8150
22	44	38.3	488	4	US-09-252-991A-28535
23	44	38.3	502	4	US-09-252-991A-31240
24	44	38.3	538	4	US-09-252-991A-21622
25	44	38.3	689	4	US-09-252-991A-32669
26	44	38.3	731	4	US-09-252-991A-17180
27	44	38.3	1124	4	US-09-252-991A-26810

28	43	37.4	202	4	US-09-252-991A-25553	Sequence 25553, A
29	43	37.4	275	4	US-09-252-991A-29227	Sequence 29227, A
30	43	37.4	284	4	US-09-252-991A-32780	Sequence 32780, A
31	43	37.4	327	2	US-08-739-485-10	Sequence 10, Appl
32	43	37.4	391	3	US-09-076-359-2	Sequence 2, Appli
33	43	37.4	411	4	US-09-543-681A-7017	Sequence 7017, Ap
34	43	37.4	533	4	US-09-252-991A-23560	Sequence 23560, A
35	43	37.4	1315	4	US-09-252-991A-22746	Sequence 22746, A
36	42.5	37.0	492	4	US-09-252-991A-28339	Sequence 28339, A
37	42.5	37.0	673	4	US-09-252-991A-26458	Sequence 26458, A
38	42	36.5	208	4	US-09-252-991A-27430	Sequence 27430, A
39	42	36.5	218	4	US-09-252-991A-17692	Sequence 17692, A
40	42	36.5	272	4	US-09-252-991A-20279	Sequence 20279, A
41	42	36.5	304	4	US-09-252-991A-17069	Sequence 17069, A
42	42	36.5	331	4	US-09-252-991A-25035	Sequence 25035, A
43	42	36.5	384	4	US-09-252-991A-26093	Sequence 26093, A
44	42	36.5	536	4	US-09-252-991A-16754	Sequence 16754, A
45	42	36.5	619	4	US-09-252-991A-21585	Sequence 21585, A

ALIGNMENTS

RESULT 1
US-08-405-175A-8
; Sequence 8, Application US/08405175A
; Patent No. 5885772
; GENERAL INFORMATION:
; APPLICANT: Aderem, Alan A.
; APPLICANT: Chen, Jianmin
; APPLICANT: Chang, Sandy
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/405.175A
FILING DATE: 16-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-121A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 309 amino acids
TYPE: amino acid
STRADEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: predicted primary structure of rat MARCKS
HYPOTHETICAL: NO
US-08-405-175A-8

Query Match 100.0%; Score 115; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 3.1e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GAQFSKTAAGGAAERPGEAAVA 24

Db 2 GAQFSKTAAGGAAARPGGAAVA 25
|||||

RESULT 2
US-08-405-175A-5
; Sequence 5, Application US/08405175A
; Patent No. 5885772
; GENERAL INFORMATION:
; APPLICANT: Aderem, Alan A.
; APPLICANT: Chen, Jianmin
; APPLICANT: Chang, Sandy
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 16-MAR-1995
; APPLICATION NUMBER: US/08/405,175A
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-121A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: Predicted primary structure of human MARCKS
; HYPOTHETICAL: NO
US-08-405-175A-5

Query Match 100.0%; Score 115; DB 2; Length 332;
Best Local Similarity 100.0%; Pred No. 3.3e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAQFSKTAAGGAAARPGGAAVA 24
|||||

Db 2 GAQFSKTAAGGAAARPGGAAVA 25
|||||

RESULT 3
US-08-405-175A-7
; Sequence 7, Application US/08405175A
; Patent No. 5885772
; GENERAL INFORMATION:
; APPLICANT: Aderem, Alan A.
; APPLICANT: Chen, Jianmin
; APPLICANT: Chang, Sandy
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey

; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/405,175A
; FILING DATE: 16-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-121A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: Predicted primary structure of murine MARCKS
; HYPOTHETICAL: NO
US-08-405-175A-7

Query Match 96.5%; Score 111; DB 2; Length 309;
Best Local Similarity 95.8%; Pred. No. 1.3e-09;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAQFSKTAAGGAAARPGGAAVA 24
|||||

Db 2 GAQFSKTAAGGAAARPGGAAVA 25
|||||

RESULT 4
US-08-405-175A-6
; Sequence 6, Application US/08405175A
; Patent No. 5885772
; GENERAL INFORMATION:
; APPLICANT: Aderem, Alan A.
; APPLICANT: Chen, Jianmin
; APPLICANT: Chang, Sandy
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/405,175A
; FILING DATE: 16-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-121A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521

;; FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A

```

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26728
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26728

Query Match
Best Local Similarity 42.6%; Score 49; DB 4; Length 314;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 10 KGEAAARPGGAAY 23
Db 30 KGQAAVEQPGQALV 43

RESULT 9
US-09-252-991A-31571
; Sequence 31571, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31571
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31571

Query Match
Best Local Similarity 41.7%; Score 48; DB 4; Length 408;
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 KTAAKGEAAARPGGAAYA 24
Db 302 RAAGLGAAARPGGTAPAA 320

RESULT 10
US-09-252-991A-28596
; Sequence 28596, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28596
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28596

Query Match
Best Local Similarity 40.9%; Score 47; DB 4; Length 996;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 9 AKGEAAARPGGAAYA 24
Db 696 ASGEGFAERPGRNAAA 711

RESULT 11
US-09-252-991A-17290
; Sequence 17290, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17290
; LENGTH: 1257
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17290

Query Match
Best Local Similarity 40.9%; Score 47; DB 4; Length 1257;
Matches 12; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 1 GAQFSKTAAKGEAAARPGGAAYA 24
Db 485 GEEHLAIAAVGEEAARVGERLAA 508

RESULT 12
US-09-252-991A-20341
; Sequence 20341, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20341
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20341

Query Match
Best Local Similarity 40.4%; Score 46.5; DB 4; Length 416;
Matches 11; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY 1 GAQFSKTAAKGEAAARPGGE 20
Db 117 GAQVAQAALARGAARREGE 137

RESULT 13
US-09-252-991A-28596
; Sequence 28596, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28596
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28596
```

```
US-09-252-991A-28823
; Sequence 28823, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28823
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28823

Query Match          40.0%; Score 46; DB 4; Length 207;
Best Local Similarity 45.5%; Pred. No. 11;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGFAAERPGEAA 22
   ||| : : : |||
Db 80 GAQAPGRCRGATGADRPGGAA 101

RESULT 14
US-09-252-991A-17603
; Sequence 17603, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17603
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17603

Query Match          40.0%; Score 46; DB 4; Length 380;
Best Local Similarity 64.7%; Pred. No. 21;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 8 AAKGEAAERPGEAAVA 24
   ||| : |||
Db 145 AAVGQARAERAAVA 161

RESULT 15
US-09-252-991A-29974
; Sequence 29974, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
```

```
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29974
; LENGTH: 758
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29974

Query Match          40.0%; Score 46; DB 4; Length 758;
Best Local Similarity 56.2%; Pred. No. 45;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 9 AKGEAAERPGEAAVA 24
   ||| : |||
Db 701 AHGESAADQGERLAA 716

Search completed: October 7, 2004, 17:36:34
Job time : 16 secs
```

This Page Blank (uspto)

F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-98/Product: outer membrane protein H.8 #status predicted <MAT>

Query Match 44.3%; Score 51; DB 2; Length 88;
Best Local Similarity 57.1%; Pred. No. 4.3;
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 APAAEGAGAEVKRASAERAKQA 23
DB 46 APAAEAAAEAPAAEAPAAEA 66

RESULT 3

T33110

hypothetical protein C18H7.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000

C;Accession: T33110

R;Tin-Wollam, A.; Fronick, W.

submitted to the EMBL Data Library, May 1998

A;Description: The sequence of C. elegans cosmid C18H7.

A;Reference number: Z31284

A;Accession: T33110

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-460 <TIN>

A;Cross-references: EMBL:AF067607; PIDN:AACT17641.1; GSPDB:GN000022; CESP:C18H7.3

A;Experimental source: strain Bristol N2; clone C18H7

C;Genetics:

A;Gene: CESP:C18H7.3

A;Map position: 4

A;Introns: 84/1

C;Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match 44.3%; Score 51; DB 2; Length 460;
Best Local Similarity 57.1%; Pred. No. 19;
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 APAAEGAGAEVKRASAERAKQA 23

DB 308 APAAEGAGGAEPPAGAAAPDA 328

RESULT 4

DB2386

methyl-accepting chemotaxis protein VCA1034 [imported] - Vibrio cholerae (strain N16961

C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C;Accession: D82386

R;Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.

l, R.K.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: AB2035; MUID:20406833; PMID:10952301

A;Accession: D82386

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-666 <HEI>

A;Cross-references: GB:AE004429; GB:AE003853; NID:g9658474; PIDN:AAF96929.1; GSPDB:GN001

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VCA1034

A;Map position: 2

Query Match 44.3%; Score 51; DB 2; Length 666;
Best Local Similarity 60.0%; Pred. No. 26;
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 APAAEGAGAEVKRASAERAKQA 22

DB 441 ARSTGAKNDEVRADAEAKQ 460

RESULT 5

S42571

subtilisin inhibitor-like peptide 2 - Streptomyces rochei

N;Alternate names: SII protein 2

C;Species: Streptomyces rochei

C;Date: 07-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999

C;Accession: S42571; PC1266

R;Taguchi, S.; Kojima, S.; Terabe, M.; Miura, K.; Momose, H.

Eur. J. Biochem. 220, 911-918, 1994

A;Title: Comparative studies on the primary structures and inhibitory properties of subtil

A;Reference number: S42571; MUID:94192683; PMID:8143745

A;Accession: S42571

A;Molecule type: protein

A;Residues: 1-113 <TAG>

A;Note: the source is designated as Streptomyces parvulus

R;Taguchi, S.; Kikuchi, H.; Kojima, S.; Kumagai, I.; Nakase, T.; Miura, K.; Momose, H.

Biosci. Biotechnol. Biochem. 57, 522-524, 1993

A;Title: High frequency of SSI-like protease inhibitors among Streptomyces.

A;Reference number: PC1260; MUID:93222542; PMID:7763545

A;Accession: PC1266

A;Molecule type: protein

A;Residues: 1-39 <TA2>

A;Experimental source: strain 2283

C;Superfamily: plasminostreptin

Query Match 43.5%; Score 50; DB 2; Length 113;
Best Local Similarity 52.6%; Pred. No. 7.3;
Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 CTAPAAEGAGAEVKRASAER 19

DB 41 GTHPADAACAELRAGGD 59

RESULT 6

I41082

S-carboxymethyl-2-hydroxyumconate semialdehyde dehydrogenase (EC 1.2.1.-) - Escherichia

C;Species: Escherichia coli

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C;Accession: I41082; A61080

R;Roper, D.I.; Stringfellow, J.M.; Cooper, R.A.

Gene 156, 47-51, 1995

A;Title: Sequence of the hpcC and hpcG genes of the meta-fission homoproteocatechuic acid

A;Reference number: I41082; MUID:95255666; PMID:7737515

A;Accession: I41082

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-468 <REG>

A;Cross-references: EMBL:X81322; NID:g587109; PIDN:CAA57102.1; PID:g587110

R;Fawcett, T.; Garrido-Perterra, A.; Cooper, R.A.

FEMS Microbiol. Lett. 57, 307-312, 1989

A;Title: 5-carboxymethyl-2-hydroxyumconic semialdehyde dehydrogenases of Escherichia coli

A;Reference number: A61080

A;Accession: A61080

A;Molecule type: protein

A;Residues: 1-19, 'Q', 21-34 <FAM>

A;Experimental source: strain C

C;Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology

C;Keywords: oxidoreductase

F;40-298/Domain: aldehyde dehydrogenase homology <ALDD>

F;244,278/Active site: Glu, Cys #status predicted

Query Match 43.5%; Score 50; DB 1; Length 468;
Best Local Similarity 50.0%; Pred. No. 26;
Matches 11; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 APAAEGAGAEVKRASAERAKQA 24

DB 31 ADVASGGAEINQAVATAKEAF 52

```
RESULT 7
T13161
N:Kinase anchor protein 95 - human
N:Alternate names: protein DKFZp586B1222.1
C:Species: Homo sapiens (man)
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C:Accession: T13161; T08786
R:Lamerdin, J.E.; McCready, P.M.; Skowronski, E.; Viswanathan, V.; Burkhardt-Schultz, K.;
J.; Dangnan, L.; Erler, A.; Christensen, M.; Georgescu, A.; Avila, J.; Liu, S.; Attix
J.; Quan, G.; Krommiller, B.; Arellano, A.; Saunders, C.; Ow, D.; Nolan, M.; Truong, S.
submitted to the EMBL Data Library, October 1998
A:Authors: Kobayashi, A.; Olsen, A.S.; Carrano, A.V.
A:Description: Sequence analysis of a 1.5 Mb OLFIR-rich region in 19p13.1.
A:Reference number: Z17617
A:Accession: T13161
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-692 <LAM>
A:Cross-references: EMBL:AC005785; NID:g3702290; PID:g3702291; PIDN:AAC62838.1
R:Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, May 1999
A:Reference number: Z16468
A:Accession: T08786
A:Molecule type: mRNA
A:Residues: 335-692 <KOE>
A:Cross-references: EMBL:AL050160
A:Experimental source: adult uterus; clone DKFZp586B1222
C:Genetics:
A:Map position: 19
A:Introns: 7/1; 20/1; 31/1; 124/2; 287/3; 331/1; 346/3; 358/1; 387/2; 434/3; 466/1; 509/1
A:Note: DKFZp586B1222.1

Query Match 43.5%; Score 50; DB 2; Length 692;
Best Local Similarity 52.4%; Pred. No. 36;
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 APAAGAGAEVVRASAEAKQA 23
||| ||| ||| ||| |||
Db 668 APAAPAAADAEEVQTDAESKDA 688

RESULT 8
S21023
N:photosystem II protein psbt - upland cotton
N:Alternate names: PSII-T protein
C:Species: Gossypium hirsutum (upland cotton)
C>Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 21-Jul-2000
C:Accession: A57500; S21023
R:Kapazoglou, A.; Sagliocco, F.; Dure III, L.
J. Biol. Chem. 270, 12197-12202, 1995
A:Title: PSII-T, a new nuclear encoded luminal protein from photosystem II. Targeting ar
A:Reference number: A57500; MUID:95263572; PMID:7744870
A:Accession: A57500
A:Molecule type: DNA
A:Residues: 1-105 <KAP>
A:Cross-references: GB:X54092; NID:g18511; PIDN:CAA38027.1; PID:g18512
A:Note: submitted to the EMBL Data Library, July 1990
C:Keywords: transmembrane protein

Query Match 42.6%; Score 49; DB 2; Length 105;
Best Local Similarity 68.8%; Pred. No. 9.4;
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 9 AGAEVVRASAEAKQAF 24
||| ||| ||| ||| |||
Db 75 ATAEPKRGSAEAKKAY 90

RESULT 9
T04347
N:teosinte branched1 protein - maize (fragment)
C:Species: Zea mays (maize)
```

```
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: T04347
R:Doebley, J.; Stec, A.O.; Hubbard, L.
Nature 386, 485-488, 1997
A:Title: The evolution of apical dominance in maize.
A:Reference number: Z15300; MUID:97242406; PMID:9087405
A:Accession: T04347
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-382 <DOE>
A:Cross-references: EMBL:U94494; NID:g2051378; PIDN:AAB53060.1; PID:g2051979
A:Experimental source: strain Pioneer Inbred AP9
C:Genetics:
A:Gene: tb1
A:Map position: 1.09

Query Match 42.6%; Score 49; DB 2; Length 382;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 APAAGAGAEVVRASAEAKQ 22
||| ||| ||| ||| |||
Db 104 ASGAEAGSASLDRAAAARK 123

RESULT 10
E83112
30S ribosomal protein S7 PA4267 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 12-Jun-2003
C:Accession: E83112
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bri
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathog
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: E83112
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-156 <STO>
A:Cross-references: GB:AE004842; GB:AE004091; NID:g9950481; PIDN:AAG07655.1; GSPDB:GN0013
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: rpsG; PA4267
C:Superfamily: ribosomal protein S7

Query Match 42.2%; Score 48.5; DB 2; Length 156;
Best Local Similarity 52.0%; Pred. No. 16;
Matches 13; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

QY 5 AAGGAGAEVVRAS-----AAKQAF 24
||| ||| ||| ||| |||
Db 127 AAEGRGAAGVKKREDVHRMAEANKAF 151

RESULT 11
S26826
N:histone H1 - maize
C:Species: Zea mays (maize)
C>Date: 12-Nov-1997 #sequence_revision 05-Dec-1997 #text_change 23-Jul-1999
C:Accession: S26826
R:Pazafimahatratra, P.; Chaubet, N.; Philipps, G.; Gigot, C.
Nucleic Acids Res. 19, 1491-1496, 1991
A:Title: Nucleotide sequence and expression of a maize H1 histone cDNA.
A:Reference number: S26826; MUID:91227140; PMID:1709276
A:Accession: S26826
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-246 <RAZ>
A:Cross-references: EMBL:X57077; NID:g22320; PIDN:CAA40362.1; PID:g22321
A:Note: the authors did not translate the codon for residue 1
C:Superfamily: histone H1
```

ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:32:03 ; Search time 8 Seconds
(without alignments)
156,210 Million cell updates/sec

Title: US-10-802-644-2

Perfect score: 115

Sequence: 1 GTAPAAEGAGAEVKRAAEAKQAF 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query %	ID	Description
1	52	45.2	1076	YE38 SCHPO	O13992 schizosacch
2	51	44.3	88	H82 NIGO	P11910 neisseria g
3	50	43.5	113	SS12 STRO	P29607 streptomyce
4	50	43.5	468	HPCC ECOLI	P42269 escherichia
5	50	43.5	692	AKA8 HUMAN	O43823 homo sapien
6	49	42.6	105	PS22 GOSHI	P31336 gossypium h
7	49	42.6	518	TPM4 DROME	P49455 drosophila
8	48.5	42.2	156	RS7 PSEAE	Q9HWD1 pseudomonas
9	48.5	42.2	156	RS7 PSEPK	Q88GN9 pseudomonas
10	48.5	42.2	156	RS7 PSESM	Q889X5 pseudomonas
11	48	41.7	245	H1 MAIZE	P23444 zea mays (m
12	47.5	41.3	155	RL5 CORGL	Q8NT20 corynebacte
13	47	40.9	110	RLA3 SCHPO	P17477 schizosacch
14	47	40.9	378	CC37 HUMAN	Q16543 homo sapien
15	47	40.9	617	41 BOVIN	Q9N179 bos taurus
16	46	40.0	110	RLA2 CRYST	O61463 cryptochic
17	46	40.0	136	RL28 HUMAN	P46779 homo sapien
18	46	40.0	136	RL28 MOUSE	P41105 mus musculu
19	46	40.0	211	NUS8 GLOVI	Q7NKO0 gloeobacter
20	46	40.0	308	RS2 STRAW	P09788 streptomyce
21	46	40.0	520	DH4C PSEPU	Q82JX7 streptomyce
22	46	40.0	637	CL16 RABIT	Q9N2G5 pseudomonas
23	45.5	39.6	110	RLA4 SCHPO	P17478 schizosacch
24	45.5	39.6	111	RLA2 ARTSA	P02399 artemia sal
25	45	39.1	191	YO76 VIBCH	Q9KPY7 vibrio chol
26	45	39.1	215	ORN MYCTU	O06174 mycobacteri
27	45	39.1	220	VG32 BPMLS	Q05241 mycobacteri
28	45	39.1	310	RS2 STRCO	O31212 streptomyce
29	45	39.1	2567	M16B HUMAN	Q8UGU5 homo sapien
30	44.5	38.7	155	RS7 MYCLE	P30764 mycobacteri
31	44.5	38.7	155	RS7 MYCSM	P41193 mycobacteri
32	44.5	38.7	155	RS7 MYCTU	P41194 mycobacteri
33	44.5	38.7	156	RS7 MYCBO	Q53539 mycobacteri

34	44.5	38.7	156	1	RS7 THICU	O50564 thiobacillu
35	44	38.3	109	1	RLAI TRYCR	P26843 trypanosoma
36	44	38.3	114	1	RLAI RAT	P19944 rattus norv
37	44	38.3	343	1	DHAS CAMJE	Q59291 campylobact
38	44	38.3	372	1	TOLA HAEIN	P44678 haemophilus
39	44	38.3	430	1	CIWC RAT	Q9ers1 rattus norv
40	44	38.3	456	1	GUNA MICBI	P26414 microbispor
41	44	38.3	467	1	HHCM HUMAN	Q05877 homo sapien
42	44	38.3	469	1	LEUZ PHOLL	Q7N127 photorhabdu
43	44	38.3	507	1	LATI HUMAN	Q01650 homo sapien
44	43.5	37.8	452	1	NIFB RHOCA	P17434 rhodobacter
45	43	37.4	107	1	SSI3 STRCO	P29608 streptomyce

ALIGNMENTS

RESULT 1
YE38 SCHPO STANDARD; PRT; 1076 AA.
AC O13992; O36013;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C20G4.08 on chromosome I.
GN SPAC20G4.08 OR SPAC4F10.01.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; 298600; CAB11257.1; -;
DR EMBL; 298980; CAB11705.1; -;
DR PIR; T38122; T38122.

```

DR   GeneDB SPombe; SPAC20G4.08; -.
KW   Hypothetical protein.
SQ   SEQUENCE 1076 AA; 119230 MW; E8FDDC3630951430 CRC64;

Query Match 45.2%; Score 52; DB 1; Length 1076;
Best Local Similarity 39.1%; Pred. No. 18;
Matches 9; Conservative 7; Mismatches 0; Gaps 0;

Qy 1 GTAPAAEGAGAEVKGASAEAKQA 23
Db 95 GAKPSGTASGADVXRSDSESTE 117

RESULT 2
H82_NEIGO
ID H82_NEIGO STANDARD; PRT; 88 AA.
AC P11910;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Outer membrane protein H.8 precursor.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 700625 / FA 1090;
RA MEDLINE=89237899; PubMed=2497298;
RX Woods J.P., Spinola S.M., Strobel S.M., Cannon J.G.;
RT "Conserved lipoprotein H.8 of pathogenic Neisseria consists entirely
of pentapeptide repeats.";
RL Mol. Microbiol. 3:43-48(1989).
CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
anchor.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X12627; CAA311145.1; -.
DR PIR; S02720; S02720.
DR InterPro; IPR000437; Prok_lipoprot_S.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Repeat; Outer membrane; Lipoprotein; Signal; Palmitate.
FT SIGNAL 1 17
FT CHAIN 18 88 OUTER MEMBRANE PROTEIN H.8.
FT LIPID 18 18 N-palmitoyl cysteine.
FT LIPID 18 18 S-diacylglycerol cysteine.
FT DOMAIN 23 87 13 X 5 AA TANDEN REPEATS OF A-A-E-A-P.
SQ SEQUENCE 88 AA; 8023 MW; 834DDFAE049FCC21 CRC64;

Query Match 44.3%; Score 51; DB 1; Length 88;
Best Local Similarity 47.1%; Pred. No. 25;
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 3 APAAEGAGAEVKGASAEAKQA 23
Db 46 APAAEAAAAEAPAAEAPAAE 66

RESULT 3
SS12_STRRO
ID SS12_STRRO STANDARD; PRT; 113 AA.
AC P29607; Q9R2G7;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Subtilisin inhibitor-like protein-2 (SIL-2) (SIL2).
```

```

OS Streptomyces rochei (Streptomyces parvullus).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1928;
RN [1]
RP SEQUENCE.
RC STRAIN=KCC 2283;
RX MEDLINE=94192683; PubMed=8143745;
RA Taguchi S., Kojima S., Terabe M., Miura K.-I., Momose H.;
RT "Comparative studies on the primary structures and inhibitory
properties of subtilisin-trypsin inhibitors from Streptomyces.";
RL Eur. J. Biochem. 220:911-918(1994).
RN [2]
RP SEQUENCE OF 1-39.
RX MEDLINE=93146392; PubMed=1490613;
RA Taguchi S., Kojima S., Kumagai I., Ogawara H., Miura K.-I., Momose H.;
RT "Isolation and partial characterization of SSI-like protease
inhibitors from Streptomyces.";
RL FEMS Microbiol. Lett. 78:293-297(1992).
RN [3]
RP SEQUENCE OF 1-39.
RC STRAIN=KCC 2283;
RX MEDLINE=93222542; PubMed=7763545;
RA Taguchi S., Kikuchi H., Kojima S., Kumagai I., Nakase T., Miura K.-I.,
RA Momose H.;
RT "High frequency of SSI-like protease inhibitors among Streptomyces.";
RL Biosci. Biotechnol. Biochem. 57:522-524(1993).
CC -!- FUNCTION: INHIBITOR OF SUBTILISIN BPN' AND TRYPSIN.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the serine protease inhibitors SSI family.
DR PIR; S42571; S42571.
DR HSP; P01006; 3SSI.
DR InterPro; IPR000691; Strep_subst_inhib.
DR Pfam; PF00720; SSI; 1.
DR PRINTS; PR00294; SSBLNINHTR.
DR ProDom; PD004028; Strep_subst_inhib; 1.
DR PROSITE; PS00999; SSI; 1.
KW Serine protease inhibitor.
FT DISULFID 35 50 BY SIMILARITY.
FT DISULFID 71 101 BY SIMILARITY.
FT ACT_SITE 73 74 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 113 AA; 11568 MW; B85B6D10787D40FC CRC64;

Query Match 43.5%; Score 50; DB 1; Length 113;
Best Local Similarity 52.6%; Pred. No. 4.2;
Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GTAPAAEGAGAEVKGASAE 19
Db 41 GTHPAADAACAEELRRAGGD 59

RESULT 4
HPCC_ECOLI
ID HPCC_ECOLI STANDARD; PRT; 468 AA.
AC P42269;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 5-carboxymethyl-2-hydroxymuconate semialdehyde dehydrogenase
(DE 1.2.1.-) (CHMS dehydrogenase).
GN HPCC.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C;
RX MEDLINE=95255666; PubMed=7737515;
RA Roper D.I., Stringfellow J.M., Cooper R.A.;
RT "Sequence of the hpcC and hpcG genes of the meta-fission
```

RT homoprotocatechuic acid pathway of *Escherichia coli* C: nearly 40%
 RT amino-acid identity with the analogous enzymes of the catechol
 RT pathway";
 RL Gene 156:47-51(1995).

CC -!- FUNCTION: TRANSFORMS 5-CARBOXYMETHYL-2-HYDROXY-MUCONIC
 CC SEMIALDEHYDE (CHMS) INTO 5-CARBOXYMETHYL-2-HYDROXY-MUCONIC ACID
 CC (CHM).

CC -!- PATHWAY: Homoprotocatechuate catabolism; second step.
 CC -!- SIMILARITY: Belongs to the aldehyde dehydrogenase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X81322; CAA57102.1; -.
 CC PIR; I41082; I41082.
 CC HSP; P05091; 1CW3.
 CC InterPro; IPR002086; Aldehyde_dehydr.

CC PROSITE; PS00070; ALDEHYDE_DEHYDR_GLY; 1.
 CC PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
 CC Aromatic hydrocarbons catabolism; Oxidoreductase.

FT ACT SITE 244 244 BY SIMILARITY.
 FT ACT SITE 278 278 BY SIMILARITY.

CC SEQUENCE 468 AA; 50827 MW; 93133200D2FDB77 CRC64;

Query Match 43.5%; Score 50; DB 1; Length 468;
 Best Local Similarity 50.0%; Pred. No. 16;
 Matches 11; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

CC 3 APAAGAGAEVKEASAEAKQAF 24
 CC 31 ADVASGGAEINQAVATAKEAF 52

RESULT 5
 AKAS HUMAN
 ID AKAS HUMAN STANDARD; PRT; 692 AA.
 AC O43823;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE A-kinase anchor protein 8 (A-kinase anchor protein 95 kDa) (AKAP 95).
 GN AKAP8 OR AKAP95.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cerebellum, and Testis;
 RX MEDLINE=98142017; PubMed=9473338;
 RA Eide T., Coghlan V., Oerstavik S., Holsve C., Solberg R.,
 RA Skaelhegg B.S., Lamb N.J.C., Langeberg L., Fernandez A., Scott J.D.,
 RA Jahnson T., Tasken K.;
 RT "Molecular cloning, chromosomal localization, and cell cycle-dependent
 RT subcellular distribution of the A-kinase anchoring protein, AKAP95";
 RL Exp. Cell Res. 238:305-316(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
 RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Scilwagen S.,
 RA Phan H., Velasco N., Do L., Regala W., Terry A., Barnes J.,
 RA Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,
 RA Attix C., Andreise T., Frankheim M., Amico-Keller G., Coefield J.,
 RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,
 RA Arellano A., Saunders C., Ow D., Nolan M., Trong S., Kobayashi A.,
 RA Olsen A.S., Carrano A.V.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Anchoring protein that mediates the subcellular
 CC compartmentation of CAMP-dependent protein kinase (PKA type II).
 CC -!- SUBUNIT: Binds to dimeric RII-alpha regulatory subunit of PKA
 CC during mitosis.

CC -!- SUBCELLULAR LOCATION: Nuclear. Associated with the nuclear matrix.
 CC Redistributed and detached from condensed chromatin during
 CC mitosis.

CC -!- TISSUE SPECIFICITY: Highly expressed in heart, liver, skeletal
 CC muscle, kidney and pancreas.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; Y11997; CAA72722.1; -.
 CC EMBL; AC005785; AAC62838.1; -.
 CC PIR; T13161; T13161.
 CC Genew; HGNC:378; AKAP8.
 CC GK; O43823; -.

CC MIM; 604692; -.
 CC GO; GO:0005634; C:nucleus; TAS.
 CC GO; GO:0007067; P:mitosis; TAS.
 CC GO; GO:0007165; P:signal transduction; TAS.
 CC InterPro; IPR007071; AKAP95.
 CC InterPro; IPR007087; Znf_C2H2.
 CC Pfam; PF04988; AKAP95; 1.
 CC SMART; SM00355; Znf_C2H2; 1.

CC Nuclear protein; Zinc-finger; DNA-binding.
 FT DOMAIN 107 118 POLY-GLY
 FT DOMAIN 368 377 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT ZN FING 392 414 C2H2-TYPE.
 FT ZN FING 481 504 C2H2-TYPE.
 FT DOMAIN 572 589 RII-BINDING (BY SIMILARITY).
 CC SEQUENCE 692 AA; 76108 MW; CBGD5F014FD94B66 CRC64;

Query Match 43.5%; Score 50; DB 1; Length 692;
 Best Local Similarity 52.4%; Pred. No. 22;
 Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

CC 3 APAAGAGAEVKEASAEAKQAF 23
 CC 668 APAPAAADAEVEQTDAESKDA 688

RESULT 6
 PST2 GOSHI
 ID PST2 GOSHI STANDARD; PRT; 105 AA.
 AC P11336;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Photosystem II 5 kDa protein, chloroplast precursor (PSII-T) (Light-
 DE regulated unknown 11 kDa protein).
 GN PSBT.
 OS Gossypium hirsutum (Upland cotton).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 OX NCBI_TaxID=3635;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Coker 201; TISSUE=Leaf;
 RX MEDLINE=95263572; PubMed=7744870;
 RA Kapazoglou A., Sagliocco F., Dure L. III;
 RT "PSII-T, a new nuclear encoded luminal protein from photosystem II.
 RT Targeting and processing in isolated chloroplasts";
 RL J. Biol. Chem. 270:12197-12202(1995).
 CC -!- FUNCTION: May be a component of the oxygen-evolving complex.
 CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated

```
CC      with the photosystem II complex.
CC      -!- PWM: The maturation of the PSII-T precursor to its final form
CC      occurs through a two step process. First, a stromal intermediate
CC      is formed, which, upon translocation into the thylakoid membrane,
CC      is processed to the mature protein.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL; X54092; CAA38027.1; -.
CC      PIR; A57500; S21023.
CC      Photosynthesis; Photosystem II; Chloroplast; Transit peptide;
CC      Thylakoid; Membrane.
CC      TRANSIT 1 77 CHLOROPLAST.
CC      CHAIN 78 105 PHOTOSYSTEM II 5 kDa PROTEIN.
CC      DOMAIN 62 67 POLY-ALA.
CC      SEQUENCE 105 AA; 11009 MW; D0C9DA37541213DE CRC64;
CC
CC      Query Match 42.6%; Score 49; DB 1; Length 105;
CC      Best Local Similarity 68.8%; Pred. No. 5.4;
CC      Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
CC
CC      QY 9 AGAEVKRASAEAKQAF 24
CC      DB 75 ATAEPKRGSAEAKKAY 90
CC
CC      RESULT 7
CC      ID TPW4 DROME STANDARD; PRT; 518 AA.
CC      AC P49455; P49456; Q24425; Q24426;
CC      DT 01-FEB-1996 (Rel. 33, Created)
CC      DT 01-FEB-1996 (Rel. 33, Last sequence update)
CC      DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC      DE Tropomyosin I, isoforms 33/34 (Tropomyosin II).
CC      TM1 OR TM11.
CC      OS Drosophila melanogaster (Fruit fly).
CC      OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
CC      OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
CC      OC Ephydroidea; Drosophilidae; Drosophila.
CC      OX NCBI_TaxID=7227;
CC      RN [1]
CC      RP SEQUENCE FROM N.A. (ISOFORMS 33 AND 34).
CC      RC TISSUE=Embryo, and Pupae;
CC      RX MEDLINE=89127197; PubMed=2851721;
CC      RA Hanke P.D., Storti R.V.;
CC      RT "The Drosophila melanogaster tropomyosin II gene produces multiple
CC      RT proteins by use of alternative tissue-specific promoters and
CC      RT alternative splicing."
CC      RL Mol. Cell. Biol. 8:3591-3602(1988).
CC      RN [2]
CC      RP SEQUENCE FROM N.A. (ISOFORMS 33 AND 34).
CC      RC STRAIN=Oregon-R; TISSUE=Pupae;
CC      RX MEDLINE=87064486; PubMed=3097506;
CC      RA Karlik C.C., Fyrberg E.A.;
CC      RT "Two Drosophila melanogaster tropomyosin genes: structural and
CC      RT functional aspects."
CC      RL Mol. Cell. Biol. 6:1965-1973(1986).
CC      CC -!- ALTERNATIVE PRODUCTS:
CC      Event=Alternative splicing; Named isoforms=5;
CC      Comment=Additional isoforms seem to exist;
CC      Name=33; Synonyms=9C;
CC      IsoId=P49455-1; Sequence=Displayed;
CC      Name=Muscle; Synonyms=9D;
CC      IsoId=P06754-1; Sequence=External;
CC      Name=Non-muscle; Synonyms=Cytoskeletal;
CC      IsoId=P06754-2; Sequence=External;
CC      Name=9A;
CC
CC      IsoId=P06754-3; Sequence=External;
CC      Name=34; Synonyms=9B;
CC      IsoId=P49455-2; Sequence=VSP 006623, VSP 006624, VSP 006625;
CC      -!- TISSUE SPECIFICITY: Both isoforms are only expressed in indirect
CC      flight muscles.
CC      -!- DEVELOPMENTAL STAGE: Both isoforms are expressed during pupal and
CC      adult stages.
CC      -!- DOMAIN: The molecule is in a coiled coil structure. The sequence
CC      exhibits a prominent seven-residues periodicity.
CC      -!- SIMILARITY: Belongs to the tropomyosin family.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL; X76208; CAA53800.1; -.
CC      EMBL; X76208; CAA53801.1; -.
CC      EMBL; K02620; AAA28967.1; ALT_SEQ.
CC      EMBL; L00355; AAA28967.1; JOINED.
CC      EMBL; L00356; AAA28967.1; JOINED.
CC      EMBL; L00357; AAA28967.1; JOINED.
CC      EMBL; L00358; AAA28967.1; JOINED.
CC      EMBL; L00359; AAA28967.1; JOINED.
CC      EMBL; L00360; AAA28967.1; JOINED.
CC      EMBL; L00362; AAA28967.1; JOINED.
CC      EMBL; M12840; AAA28967.1; JOINED.
CC      EMBL; K02621; AAA28968.1; -.
CC      EMBL; M12840; AAA28968.1; JOINED.
CC      EMBL; L00355; AAA28968.1; JOINED.
CC      EMBL; L00356; AAA28968.1; JOINED.
CC      EMBL; L00357; AAA28968.1; JOINED.
CC      EMBL; L00358; AAA28968.1; JOINED.
CC      EMBL; L00359; AAA28968.1; JOINED.
CC      EMBL; L00360; AAA28968.1; JOINED.
CC      EMBL; L00362; AAA28968.1; JOINED.
CC      FlyBase; FBgn003721; Tm1.
CC      GO; GO:004541; P:pole plasm oskar mRNA localization; IMP.
CC      InterPro; IPR000533; Tropomyosin.
CC      Pfam; PF00261; Tropomyosin; 1.
CC      PRINTS; PR00194; TROPOMYOSIN.
CC      PROSITE; PS00326; TROPOMYOSIN; 1.
CC      Muscle protein; Cytoskeleton; Actin-binding; Coiled coil;
CC      Alternative splicing; Multigene family.
CC      COILED COIL (POTENTIAL).
CC      DOMAIN 14 267
CC      DOMAIN 287 518
CC      VARSPLIC 259 293
CC
CC      VARSPLIC 300 367
CC
CC      VARSPLIC 391 518
CC
CC      CONFLICT 106 114
CC      CONFLICT 119 119
CC      CONFLICT 183 183
CC      CONFLICT 199 199
CC      CONFLICT 503 503
CC      SEQUENCE 518 AA; 54558 MW; 153D0872CF9DB6EA CRC64;
CC
```



```
Query Match          42.6%; Score 49; DB 1; Length 518;
Best Local Similarity 57.1%; Pred. No. 23;
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 APAAGAGAEVVKRAAEAKQA 23
DB 476 APAAGAAFAEAPAEATRAEA 496

RESULT 8
RS7_PSEAF
ID RS7_PSEAF STANDARD; PRT; 156 AA.
AC Q9HWD1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S7.
GN RPSG OR PA4267.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC
CC -!- FUNCTION: One of the primary rRNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the head
CC domain of the 30S subunit. Is located at the subunit interface
CC close to the decoding center, probably blocks exit of the E-site
CC tRNA (By similarity).
CC
CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Contacts proteins S9
CC and S11 (By similarity).
CC
CC -!- SIMILARITY: Belongs to the S7P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE004842; AAC07655.1; -.
CC F1R; E83112; E83112.
CC HSP; P22744; LHUS.
CC HAMAP; MF 00480; -.
CC InterPro; IPR002335; Ribosomal_S7.
CC InterPro; IPR005717; Ribosomal_S7_b/o.
CC Pfam; PF00177; Ribosomal_S7; 1.
CC ProDom; PD000817; Ribosomal_S7; 1.
CC TIGRFAMs; TIGR01029; rpsG_bact; 1.
CC PROSITE; PS00052; RPSG; 1.
CC Ribosomal protein; RNA-binding; rRNA-binding; tRNA-binding;
CC Complete proteome.
CC SEQUENCE 156 AA; 17504 MW; B747FDB710DDC14 CRC64;

Query Match          42.2%; Score 48.5; DB 1; Length 156;
Best Local Similarity 52.0%; Pred. No. 9;
Matches 13; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

QY 5 AAEAGAGAEVKRAS-----AEAKQAF 24
DB 127 AAEAGAAVKKREDVHRMAEANKAF 151

RESULT 10
RS7_PSESM
ID RS7_PSESM STANDARD; PRT; 156 AA.
AC Q88X5;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)

QY 5 AAEAGAGAEVKRAS-----AEAKQAF 24
DB 127 AAEAGAAVKKREDVHRMAEANKAF 151
```

```
RESULT 9
RS7_PSEPK
ID RS7_PSEPK STANDARD; PRT; 156 AA.
AC Q88QN9;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 30S ribosomal protein S7.
GN RPSG OR PP0450.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzez A.,
RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
CC -!- FUNCTION: One of the primary rRNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the head
CC domain of the 30S subunit. Is located at the subunit interface
CC close to the decoding center, probably blocks exit of the E-site
CC tRNA (By similarity).
CC
CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Contacts proteins S9
CC and S11 (By similarity).
CC
CC -!- SIMILARITY: Belongs to the S7P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE016775; AAN66080.1; -.
CC TIGR; PP0450; -.
CC HAMAP; MF 00480; -.
CC InterPro; IPR000235; Ribosomal_S7.
CC Pfam; PF00177; Ribosomal_S7; 1.
CC PROSITE; PS00052; RIBOSOMAL_S7; 1.
CC Ribosomal protein; RNA-binding; rRNA-binding; tRNA-binding;
CC Complete proteome.
CC SEQUENCE 156 AA; 17578 MW; 22E47BC15B4B921A CRC64;

Query Match          42.2%; Score 48.5; DB 1; Length 156;
Best Local Similarity 52.0%; Pred. No. 9;
Matches 13; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

QY 5 AAEAGAGAEVKRAS-----AEAKQAF 24
DB 127 AAEAGAAVKKREDVHRMAEANKAF 151

RESULT 10
RS7_PSESM
ID RS7_PSESM STANDARD; PRT; 156 AA.
AC Q88X5;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
```

```
DE 30S ribosomal protein S7.
GN RPSG OR PSPT00622.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RX MEDLINE=22834015; PubMed=12928499;
RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Gwinn M.L., Dodson R.J., Deboy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S., Brinkac L., Beanan M.J., Haft D.H.,
RA Nelson W.C., Davidov N., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H., Fedorova N., Tran B., Russell D., Berry K., Uterback T.,
RA Van Aken S.E., Feldblyum T.V., D'Ascenzo M., Deng W.-L., Ramos A.R.,
RA Alfano J.R., Carinhon S., Chatterjee A.K., Delaney T.P.,
RA Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X., Bender C.L.,
RA White O., Fraser C.M., Collmer A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomato DC3000.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186 (2003).
CC -!- FUNCTION: One of the primary rRNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the head
CC domain of the 30S subunit. Is located at the subunit interface
CC close to the decoding center, probably blocks exit of the E-site
CC tRNA (By similarity).
CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Contacts proteins S9
CC and S11 (By similarity).
CC -!- SIMILARITY: Belongs to the S7P family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE016858; MA054164.1; -.
CC TIGR; PSPT00622; -.
CC HAVAP; MF_00480; -.
CC InterPro; IPR000235; Ribosomal_S7.
CC Pfam; PF00177; Ribosomal_S7; 1.
CC PROSITE; PS00052; RIBOSOMAL_S7; 1.
CC Ribosomal protein; RNA-binding; rRNA-binding; tRNA-binding;
CC Complete proteome.
CC SEQUENCE 156 AA; 17665 MW; 26957E507E10F8D0 CRC64;
Query Match 42.2%; Score 48.5; DB 1; Length 156;
Best Local Similarity 52.0%; Pred. No. 9;
Matches 13; Conservative 2; Mismatches 5; Indels 5; Gaps 1;
Qy 5 AAEKGAGAEVKRAS-----AEAKQAF 24
|||||
Db 127 AAEKGAGAEVKREDVHRMAEANKAF 151
|||||

RESULT 11
HL_MAIZE STANDARD; PRT; 245 AA.
AC P23444;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Histone H1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]_
RP SEQUENCE FROM N.A.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
RC STRAIN=cv. W22;
RX MEDLINE=91227140; PubMed=1709276;
RA Razafimahatratra P., Chaubet N., Phillips G., Gigot C.;
RT "Nucleotide sequence and expression of a maize H1 histone cDNA.";
RL Nucleic Acids Res. 19:1491-1496(1991).
CC -!- FUNCTION: Histones H1 are necessary for the condensation of of
CC nucleosome chains into higher order structures.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the histone H1/H5 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X57077; CAA40362.1; -.
CC PIR; S26826; S26826.
CC MaizeDB; 25540; -.
CC InterPro; IPR005818; Histone_H1/H5.
CC InterPro; IPR005819; Histone_H5.
CC InterPro; IPR003216; Linkerhist_N.
CC Pfam; PF00538; linker histone; 1.
CC PRINTS; PR00624; HISTONEH5.
CC ProDom; PD000373; Linkerhist_N; 1.
CC SMART; SM00526; H15; 1.
CC Chromosomal protein; Nuclear protein; DNA-binding.
KW INIT_MET 0
FT DOMAIN 1 30 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 31 45 ARG/LYS-RICH (BASIC).
FT DOMAIN 46 121 GLOBULAR.
SQ SEQUENCE 245 AA; 25217 MW; D1FD73DAA74A6778 CRC64;
Query Match 41.7%; Score 48; DB 1; Length 245;
Best Local Similarity 47.6%; Pred. No. 16;
Matches 10; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
Qy 3 APAAEGAGAEVKRASAEAKQA 23
|||||
Db 25 APAADANAAKAKTAPKRA 45
|||||

RESULT 12
RS7_CORGL STANDARD; PRT; 155 AA.
AC Q8NT20;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S7.
GN RPSG OR CGL0494.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: One of the primary rRNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the head
CC domain of the 30S subunit. Is located at the subunit interface
CC close to the decoding center, probably blocks exit of the E-site
CC tRNA (By similarity).
CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Contacts proteins S9
CC and S11 (By similarity).
CC -!- SIMILARITY: Belongs to the S7P family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```


FT DOMAIN 467 617 CARBOXYL-TERMINAL (CTD).
SQ SEQUENCE 617 AA; 69257 MW; 5B49D5008AD900FB CRC64;
Query Match 40.9%; Score 47; DB 1; Length 617;
Best Local Similarity 54.5%; Pred. No. 51;
Matches 12; Conservative 2; Mismatches 6; Indels 2; Gaps 1;
QY 3 APAARGA--GAEVKRAAEAKQ 22
Db 358 SPAAREGGVPGAPVKKAQKETVQ 379

Search completed: October 7, 2004, 17:34:18
Job time : 10 secs

This Page Blank (uspto)

DT	01-OCT-2003	(TReMBLrel. 25, Created)
DT	01-OCT-2003	(TReMBLrel. 25, Last sequence update)


```

DR EMBL; AC108884; AAM01141.1; -.
DR EMBL; AC122144; AAM44895.1; -.
DR EMBL; AE017090; AAP53610.1; -.
DR Gramene; O8S5P3; -.
KW Hypothetical protein.
SQ SEQUENCE 183 AA; 19747 MW; 8E83AE03BEE44AB1 CRC64;

Query Match          44.3%; Score 51; DB 10; Length 183;
Best Local Similarity 45.5%; Pred. No. 35;
Matches 10; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 GTAPAAEGAGAEVKRASAEAKQ 22
DB 38 GDRPTAEGGAATKRCGREGER 59

RESULT 6
OBN2Q3 PRELIMINARY; PRT; 316 AA.
ID Q8N2Q3
AC Q8N2Q3;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hypothetical protein FL90054 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RA "NEO human cDNA sequencing project.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK074535; BAC11045.1; -.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00553; UBIQUITIN_2; 1.
KW Hypothetical protein.
FT NON_TER 316 316
SQ SEQUENCE 316 AA; 34030 MW; 167FC61940485600 CRC64;

Query Match          44.3%; Score 51; DB 4; Length 316;
Best Local Similarity 60.0%; Pred. No. 62;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 TAPAAEGAGAEVKRASAEAK 21
DB 16 SAAGAGAGAPAAASAEPK 35

RESULT 7
Q9GZF7 PRELIMINARY; PRT; 460 AA.
ID Q9GZF7
AC Q9GZF7;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN C18H7.3
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;

None;
"Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
Science 282:2012-2018 (1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Tin-Wollam A., Fronick W.;
RT "The sequence of C. elegans cosmid C18H7.";
RT Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF067607; AAF98609.1; -.
DR PIR; T33110; T33110.
DR WormPep; C18H7.3; CE17428.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR002486; Col_cuticle_N.
DR Pfam; PF01391; Collagen; 2.
DR Pfam; PF01484; Col_cuticle_N; 1.
KW Hypothetical protein.
SQ SEQUENCE 460 AA; 41016 MW; BF99CC80770E2202 CRC64;

Query Match          44.3%; Score 51; DB 5; Length 460;
Best Local Similarity 57.1%; Pred. No. 91;
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 APAAGAGAEVKRASAEAKQA 23
DB 308 APAAGAGAGGAGPAGAAAPDA 328

RESULT 8
Q8ZQ48 PRELIMINARY; PRT; 488 AA.
ID Q8ZQ48
AC Q8ZQ48;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE 4-hydroxyphenylacetate catabolism (EC 1.2.1.60).
GN HPAE OR STM1102.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RT LT2.";
RL Nature 413:852-856 (2001).
DR EMBL; AE008747; AAL20034.1; -.
DR GO; GO:0018480; F:5-carboxymethyl-2-hydroxy-5-oxo-2-pyridinecarboxaldehyde; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00171; aldehyd; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 488 AA; 53039 MW; 7CEBF3AB95FBA8CF CRC64;

Query Match          44.3%; Score 51; DB 16; Length 488;
Best Local Similarity 54.5%; Pred. No. 96;

```

Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 3 APAAGGAGAEVVKRASAEAK 24

Db 31 AEVASGGAEVQNQAAAEAK 52

RESULT 9

Q9HOT8
ID Q9HOT8 PRELIMINARY; PRT; 589 AA.
AC Q9HOT8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN DKFZP564L023.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Waubutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL136643; CAB66578.1; -.
DR HSSP; P02248; 1TBE.
DR InterPro; IPR006636; STIL.
DR InterPro; IPR000449; UBA domain.
DR Pfam; PF00627; UBA; 1.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00727; STIL; 4.
DR SMART; SM00165; UBA; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS50030; UBA; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 589 AA; 62489 MW; F41E0F394862FF7C CRC64;

Query Match 44.3%; Score 51; DB 4; Length 589;

Best Local Similarity 60.0%; Pred. No. 1.2e+02;

Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 2 TAPAAEGAGAEVVKRASAEAK 21

Db 16 SAAGAEGAGAPAAASAEPK 35

RESULT 10

Q9UMX0
ID Q9UMX0 PRELIMINARY; PRT; 589 AA.
AC Q9UMX0; Q9HA25;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Ubiquilin (P4IC-1).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20530664; PubMed=11076969;
RA Mah A.L., Perry G., Smith M.A., Monteiro M.J.;
RT "Identification of ubiquilin, a novel presenilin interactor that
increases presenilin protein accumulation.";
RL J. Cell Biol. 151:847-862 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Mah A.L., Monteiro M.J.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=20437281; PubMed=10983987;
RA Kleijnen M.F., Shih A.H., Zhou P., Kumar S., Soccio R.E.,
RA Kederisha N.L., Gill G., Howley P.M.;
RT "The nPLIC proteins may provide a link between the ubiquitination
machinery and the proteasome.";
RL Mol. Cell 6:409-419 (2000).
DR EMBL; AF176069; AAD49751.3; -.
DR EMBL; AF293384; AAG02473.1; -.
DR HSSP; P02248; 1TBE.
DR InterPro; IPR006636; STIL.
DR InterPro; IPR000449; UBA domain.
DR Pfam; PF00627; UBA; 1.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00727; STIL; 4.
DR SMART; SM00165; UBA; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS50030; UBA; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 589 AA; 62519 MW; 8B4756B6113B7025 CRC64;

Query Match 44.3%; Score 51; DB 4; Length 589;

Best Local Similarity 60.0%; Pred. No. 1.2e+02;

Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 2 TAPAAEGAGAEVVKRASAEAK 21

Db 16 SAAGAEGAGAPAAASAEPK 35

RESULT 11

Q9H3R4
ID Q9H3R4 PRELIMINARY; PRT; 589 AA.
AC Q9H3R4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE DA41.
GN DA41.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Hanaoka E.;
RT "human DA41.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB035275; BAB20436.1; -.
DR HSSP; P02248; 1TBE.
DR InterPro; IPR006636; STIL.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 1.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00727; STIL; 4.
DR SMART; SM00165; UBA; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS50030; UBA; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 589 AA; 62443 MW; 5E3CA56C57F61A25 CRC64;

Query Match

Best Local Similarity 60.0%; Pred. No. 1.2e+02;

Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 2 TAPAAEGAGAEVVKRASAEAK 21

Db 16 SAAGAEGAGAPAAASAEPK 35

```

RESULT 12
Q9KKR7
ID Q9KKR7 PRELIMINARY; PRT; 666 AA.
AC Q9KKR7
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Methyl-accepting chemotaxis protein.
GN VCA1034.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E1 Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483 (2000).
DR EMBL; AB004429; AAF96929.1; -.
DR PIR; D82386; D82386.
DR HSP; P02942; LQUT.
DR TIGR; VCA1034; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0006935; P:chemotaxis; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR004089; Chtaxis_trans.
DR InterPro; IPR003660; HAMP.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00283; MA; 1.
DR PROSITE; PS50111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS50885; HAMP; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 666 AA; 73398 MW; AE89ABDC756A9401 CRC64;

Query Match 44.3%; Score 51; DB 16; Length 666;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 APAEAGAGAEVKRASAEAKQ 22
Db 441 ARSTDGAKDEVDRADAERAKQ 460

RESULT 13
Q9LON7
ID Q9LON7 PRELIMINARY; PRT; 956 AA.
AC Q9LON7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein SC04254.
GN SC04254 OR SC04254.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX Seeger K.J., Harris D.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RESULT 14
Q828Y4
ID Q828Y4 PRELIMINARY; PRT; 224 AA.
AC Q828Y4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative peptidoglycan-binding protein.
GN SAV6527.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Onose T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;

```

RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism *Streptomyces avermitilis*.";
RL Nat. Biotechnol. 21:526-531(2003).

DR EMBL: AP005047; BAC74238.1; -.

DR EMBL; AF003047; AAC1238.1;
DR GO: GO:0016998; P:cell wall catabolism; IEA.

DR GO:0016338; F:Cell wall
DR InterPro: IPR002482; LysM.

DR ILCEP10; IPR002482; LYS1
DR Pfam: PF01476: LysM: 1.

```
DR PLAM; PF01476; LYSM; 1.
DR SMART: SM00257: LYSM: 1.
```

DR SMART; SM00257; LYSM; I.
KW Complete proteome.

RW	Complete procedure.	22577 MW:	2244CB5EC7AFB37E CRC64:
SO	SEQUENCE	224 AA:	

Query Match 43.9%; Score 50.5; DB 16; Length 224;

Best Local Similarity 40.6%; pred. No. 51;

Best Local Similarity	40.6%	Accuracy	97.8%
Matches	13; Conservative	Mismatches	4; Indels
		Gaps	6; Gaps

1 GTAPAAEG-----AGAEVKRASAEAKQA 23

[illegible]

RESULT 15

Q8RKT0

Q8RKT0	PRELIMINARY;	PRT;	158 AA.
ID	Q8RKT0		

AC Q8RKT0;

DT 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)

DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)

DE Putative serine-threonine

GN PKAD.

OS Frankia sp. (strain EuIK1).

OC Bacteria; Actinobacteria; Actinobacteria

OC Frankineae; Frank

OX NCBI_TaxID=47227;

RN [1] SEQUENCE FROM N A

RP SEQUENCE FROM N.A.

RA On C.J., An C.S.;
PI submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

RL Submitted (MAR-2002) to the EM
 DP FMBL: AY091786: AAM13700 1: -

DR EMBL; AY091786; AAM13700.1; - -
DR GO: GO:0016301: F:kinase activity: IEA:

UR GU; GU: (

KW	ALLIASE:	1
FT	NON TER	

FT	NON TER	I	I
SEQUENCE	158 AA:	16860 MW:	FFF0F0CDDDE579AD3 CRC64;

Query Match: 43.5%; Score 50; DB 2; Length 158;

Query Match 43.3%; Score 30;
Best Local Similarity 61.1%; Pred. No. 41;

Best Local Similarity	Conservative	Mismatches	Indels	Gaps
Matches	11;	1;	6;	0;

1 GTAPAEAGAGAEVKRASA 18
QV[illegible]

Search completed: October 7, 2004, 17:35:35

search completed: 00
Job time : 38.5 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:32:02 ; Search time 50.5 Seconds
(without alignments)
134.280 Million cell updates/sec

Title: US-10-802-644-2

Perfect score: 115

Sequence: 1 GTAPAAEGAGAEVTKASAEAKQAF 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	100.0	24	6	Abp97766 Scrambled
2	53	46.1	487	6	Abu22475 Protein e
3	53	46.1	589	4	Aab74670 Human pro
4	51	44.3	79	6	Abp79644 N. gonorr
5	51	44.3	316	4	Aam93216 Human pol
6	51	44.3	409	6	Abu17304 Protein e
7	51	44.3	412	7	Ada28111 Human NTR
8	51	44.3	422	6	Ada33525 Acinetoba
9	51	44.3	488	6	Abu47184 Protein e
10	51	44.3	589	4	Aab95205 Human pro
11	51	44.3	589	4	Aab94311 Human pro
12	51	44.3	589	4	Abu52717 Human bra
13	51	44.3	589	4	Aab62194 Human SAP
14	51	44.3	589	4	Aam94008 Human sto
15	51	44.3	589	7	Ada59377 Human pro
16	51	44.3	604	4	Abg16557 Novel hum
17	50	43.5	113	2	Aaw75885 Peptide i
18	50	43.5	113	3	Aay78865 Streptomy
19	50	43.5	116	3	Aab37425 Human sec
20	50	43.5	116	3	Aab37426 Human sec
21	49.5	43.0	74	5	Abp35189 Human dea
22	49.5	43.0	280	4	Aau61421 Propionib
23	49.5	43.0	280	6	Aam57940 Propionib
24	49	42.6	418	6	Ada34413 Acinetoba
25	49	42.6	518	4	Aau32416 Novel hum

ALIGNMENTS

RESULT 1

ABP97766

ID ABP97766 standard; peptide; 24 AA.

XX AC ABP97766;

XX DT 11-AUG-2003 (first entry)

XX DE Scrambled myristoylated N-terminal sequence (MANS) peptide.

XX KW Myristoylated N-terminal sequence; MANS; MARCKS; mucus secretion;
inflammatory mediator; inflammation; respiratory disease; asthma;
chronic bronchitis; chronic obstructive pulmonary disease; COPD;
bowel disease; irritable bowel syndrome; Crohn's disease;
ulcerative colitis; skin disease; rosacea; eczema; psoriasis; acne;
autoimmune disease; pain; arthritis; cystic fibrosis.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal myristate chain attached"

XX PN WO2003000027-A2.

XX XX 03-JAN-2003.

XX PF 26-JUN-2002; 2002WO-US022270.

XX PR 26-JUN-2001; 2001US-0300933P.

XX PA (UYNC-) UNIV NORTH CAROLINA STATE.

XX PI Martin LD, Adler KB, Li Y;

XX DR WPI; 2003-278239/27.

XX PT Method of regulating inflammation comprises administering a composition comprising a MANS peptide or an active fragment thereof.

XX PS Disclosure; Page 26; 54pp; English.

XX CC The present sequence represents a scrambled myristoylated N-terminal sequence (MANS) peptide. It is used as a control peptide, in the method of the invention. The MANS peptide is identical to the first 24 amino acids of MARCKS, and mediates insertion of MARCKS into membranes. The MANS peptide inhibits both mucus secretion and inflammatory mediators. The specification describes a method of regulating inflammation. The

PT autoimmune/inflammatory disorders such as acquired immunodeficiency
 PT syndrome, Cushing's disease, Addison's disease and cell proliferative
 PT disorders such as cancer.
 PS Claim 1; Page 93-95; 134pp; English.
 XX
 CC AAF81714 to AAF81740 encode the human proteases and protease inhibitors
 CC (PPIMs) given in AAB74668 to AAB74694. The PPIMs can have activities such
 CC as: anti-human immunodeficiency virus (HIV); antidiabetic; antithyroid;
 CC immunostimulant; immunomodulator; anti-inflammatory; immunosuppressive;
 CC nephrotropic; antigout; thyromimetic; cytostatic; antibacterial;
 CC fungicide; protozoacide; antiarteriosclerotic; antiatherosclerotic;
 CC virucide; antipsoriatic; and hepatotropic. PPIM polynucleotide and
 CC protein sequences can be used in the diagnosis, treatment and prevention
 CC of autoimmune/inflammatory disorders such as AIDS, DiGeorge's syndrome,
 CC severe combined immunodeficiency disease (SCID), Chediak-Higashi
 CC syndrome, Cushing's disease, Addison's disease, autoimmune thyroiditis,
 CC Crohn's disease, diabetes mellitus, Good pasture's syndrome, gout,
 CC Grave's diseases, Hashimoto's thyroiditis, Sjogren's syndrome, Werner's
 CC syndrome, viral, bacterial, fungal, parasitic, protozoal, and helminthic
 CC infections and cell proliferative disorder such as arteriosclerosis,
 CC atherosclerosis, cirrhosis, hepatitis, psoriasis and cancer. PPIM
 CC polynucleotide sequences can be used in somatic or germ-line gene therapy
 CC and in diagnosis of diseases. They can also be used in generating
 CC hybridisation probes useful in mapping the naturally occurring genomic
 CC sequences and in molecular biology techniques
 XX
 SQ Sequence 589 AA;

Query Match 46.1%; Score 53; DB 4; Length 589;
 Best Local Similarity 60.0%; Pred. No. 30;
 Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 TAPAEAGAGAEVVKRSAAK 21
 Db 16 SAAGAGAGAPAAASADAK 35

RESULT 4
 AABP79644
 ID AABP79644 standard; protein; 79 AA.
 XX
 AC AABP79644;
 XX
 DT 07-MAR-2003 (first entry)
 DE N. gonorrhoeae amino acid sequence SEQ ID 5818.
 XX Antibacterial; infection; vaccine; gene therapy.
 XX Neisseria gonorrhoeae.
 OS
 XX WO200279243-A2.
 PN 10-OCT-2002.
 PD
 XX 12-FEB-2002; 2002WO-IB002069.
 XX 12-FEB-2001; 2001GB-00003424.
 XX (CHTR-) CHIRON SPA.
 PA Fontana MR, Pizza M, Masignani V, Monaci E;
 PI WPI; 2003-058415/05.
 DR N-PSDB; AB240614.
 XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
 PT medicament for treating or preventing N. gonorrhoeae infection.
 PT Disclosure; Page 612; 815pp; English.

XX The present invention relates to proteins from Neisseria gonorrhoeae.

CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing N. gonorrhoeae
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records AAB76736-ABP81046 represent nucleic acid
 CC molecules of the invention
 XX
 SQ Sequence 79 AA;

Query Match 44.3%; Score 51; DB 6; Length 79;
 Best Local Similarity 57.1%; Pred. No. 7;
 Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 APAAGAGAEVVKRSAAKQA 23
 Db 37 APAAGAGAEVVKRSAAKQA 57

RESULT 5
 AAM93216
 ID AAM93216 standard; protein; 316 AA.
 XX
 AC AAM93216;
 XX
 DT 06-NOV-2001 (first entry)
 DE Human polypeptide, SEQ ID NO: 2620.
 XX
 KW Human; full length cDNA; cDNA synthesis; oligo-capping.
 XX
 OS Homo sapiens.
 XX EP1130094-A2.
 PN
 PD 05-SEP-2001.
 XX
 PF 07-JUL-2000; 2000EP-00114089.
 XX
 PR 08-JUL-1999; 93JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183765.
 XX
 PA (HELI-) HELIX RES INST.

XX Ota T, Nishikawa T, Isogai T, Hayaishi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX WPI; 2001-524255/58.
 DR N-PSDB; AAK94125.

PT 830 Primers useful for synthesizing full length cDNA clones and their use
 PT in genetic manipulation.

PS Claim 8; SEQ ID NO 2620; 1380pp + Sequence Listing; English.

XX The invention relates to primers for synthesising full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5' and 3' ends of the cDNA molecules have
 CC been determined. Primers for synthesising the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a polypeptide encoded by a full length
 CC human cDNA of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in CD-ROM
 CC format directly from EPO

XX Sequence 316 AA;

Query Match 44.3%; Score 51; DB 4; Length 316;
 Best Local Similarity 60.0%; Pred. No. 30;

Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 2 TAPAAEGAGAEVKRASAEAK 21
Db 16 SAAGAEGAGAPAAASAEFK 35

RESULT 6

ABU17304
ID ABU17304 standard; protein; 409 AA.

XX
AC ABU17304;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #2831.

DE Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Acinetobacter baumannii.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

DR WPI; 2003-029926/02.

DR N-PSDB; ACA21174.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 45228; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway;
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

CC K. pneumoniae or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 409 AA;

Query Match 44.3%; Score 51; DB 6; Length 409;

Best Local Similarity 60.0%; Pred. No. 40;

Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 APAAEGAGAEVKRASAEAKQ 22

Db 269 APSGEGATRCIKLALAEAKQ 288

RESULT 7

ADE28111

ID ADE28111 standard; protein; 412 AA.

XX ADE28111;

XX 29-JAN-2004 (first entry)

XX Human NTRAN protein - SEQ ID 16.

XX human; neurotransmission-associated protein; NTRAN; cytostatic;
KW immunomodulator; immune disorder; cancer; gene therapy.

XX Homo sapiens.

XX WO2003051902-A1.

XX 26-JUN-2003.

XX 12-DEC-2002; 2002WO-US040059.

XX 14-DEC-2001; 2001US-0340798P.

PR 18-MAR-2002; 2002US-0365645P.

PR 25-MAY-2002; 2002US-0367662P.

PR 10-MAY-2002; 2002US-0379887P.

PR 31-MAY-2002; 2002US-0384639P.

XX (INCY-) INCYTE GENOMICS INC.

XX Baughn MR, Bhatia U, Blake JJ, Burrill JD, Elliott VS;

PI Emerling EM, Forsythe LJ, Gietzen KJ, Gorvad AE, Griffin JA;

PI Hafalia AJA, Ho A, Jackson AA, Jiang X, Kable AE, Kearney L;

PI Khare R, Lee EA, Lee S, Lu DAM, Marquis JP, Lehr-Mason PM;

PI Ramkumar J, Richardson TW, Sprague WW, Tran UK, Chawla NK;

PI Warren BA, Yue H, Zheng W;

XX WPI; 2003-514037/48.

DR N-PSDB; ADE28133.

XX New human neurotransmission-associated proteins (NTRAN) polypeptide,
PT useful for preparing a composition for treating a disease associated with
PT decreased expression or overexpression of NTRAN e.g., cancer.

XX Claim 1; SEQ ID NO 16; 261pp; English.

XX The invention relates to a novel isolated human neurotransmission-
CC associated proteins (NTRAN) polypeptide. The polypeptide of the invention
CC demonstrates cytostatic and immunomodulator activities and may be useful
CC for preparing a composition for diagnosing or treating a disease or
CC condition associated with decreased expression or overexpression of
CC functional NTRAN including immune disorders or cancer, as well as during
CC gene therapy procedures. The current sequence is that of the human NTRAN
CC protein of the invention.

XX Sequence 412 AA;

SQ

Query Match 44.3%; Score 51; DB 7; Length 412;
Best Local Similarity 60.0%; Pred. No. 40;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 TAPAAEGAGAEVKRASAEAK 21
Db 16 SAAGAEGAGAPAAASAEK 35

RESULT 8
ADA33525
ID ADA33525 standard; protein; 422 AA.

XX AC ADA33525;
XX 20-NOV-2003 (first entry)
XX Acinetobacter baumannii protein #686.

XX Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
XX plant biocontrol agent.

OS Acinetobacter baumannii.

FN US6562958-B1.

XX 13-MAY-2003.

XX 04-JUN-1999; 99US-00328352.

XX 09-JUN-1998; 98US-0088701P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Breton G, Bush D;

XX WPI: 2003-576092/54.

XX N-PSDB; ADA29399.

XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
XX for diagnosing a bacterial disease, as components of antibacterial
XX vaccines, as targets for antibacterial drugs, or as biocontrol agents for
XX plants.

XX Example; SEQ ID NO 4812; 328pp; English.

XX The invention relates to isolated Acinetobacter baumannii nucleic acids.
XX The A. baumannii nucleic acids and polypeptides are useful as reagents
XX for diagnosing a bacterial disease, as components of antibacterial
XX vaccines, as targets for antibacterial drugs, to detect the presence of
XX A. baumannii and other Acinetobacter species in a sample, in screening
XX compounds for the ability to interfere with the A. baumannii life cycle
XX or to inhibit A. baumannii infection, and as biocontrol agents for
XX plants. The present sequence represents the amino acid sequence of an A.
XX baumannii protein.

XX Sequence 422 AA;

Query Match 44.3%; Score 51; DB 6; Length 422;
Best Local Similarity 60.0%; Pred. No. 41;
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 APAAEGAGAEVKRASAEAKQ 22
Db 282 APSGEGATRCIKLALAEAKQ 301

RESULT 9
ABU47184
ID ABU47184 standard; protein; 488 AA.

XX AC ABU47184;

XX

DT 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #32711.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Salmomella typhimurium.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI: 2003-029926/02.

XX N-PSDB; ACA51054.

XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 75108; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:

XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 488 AA;

Query Match 44.3%; Score 51; DB 6; Length 488;
Best Local Similarity 54.5%; Pred. No. 48;
Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 3 APAAEGAGAEVKRASAEAKQAF 24

Db 31 AEVAGGAEVNOVAARKEAF 52

RESULT 10
AAB95205
ID AAB95205 standard; protein; 589 AA.

XX AAB95205;
XX 26-JUN-2001 (first entry)
XX Human protein sequence SEQ ID NO:17306.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX Homo sapiens.

XX EPI074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.
XX 27-AUG-1999; 99JP-00300253.
XX 11-JAN-2000; 2000JP-00118776.
XX 02-MAY-2000; 2000JP-00183767.
XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX Claim 8; SEQ ID NO 17306; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95993 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX Sequence 589 AA;

XX Query Match 44.3%; Score 51; DB 4; Length 589;
XX Best Local Similarity 60.0%; Pred. No. 59;
XX Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 TAPAAAGAGAEVKRAAEAK 21
Db 16 SAAGAGAGAPAAAAAEPK 35

RESULT 11
AAB94311
ID AAB94311 standard; protein; 589 AA.

XX AAB94311;
XX 26-JUN-2001 (first entry)
XX Human protein sequence SEQ ID NO:14781.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX Homo sapiens.

XX EPI074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.
XX 27-AUG-1999; 99JP-00300253.
XX 11-JAN-2000; 2000JP-00118776.
XX 02-MAY-2000; 2000JP-00183767.
XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX Claim 8; SEQ ID NO 14781; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95993 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX Sequence 589 AA;

XX Query Match 44.3%; Score 51; DB 4; Length 589;
XX Best Local Similarity 60.0%; Pred. No. 59;
XX Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

18-OCT-1999; 99US-0159590P.

```
PR 11-JAN-2000; 2000JP-00118776.
PR 17-FEB-2000; 2000US-0183322P.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Aburatani H;
PI Kodama T, Midorikawa Y;
XX
XX WPI; 2001-570287/64.
DR N-PSDB; AAI93885.
XX
XX New Stomach cancer-associated genes, useful as markers in blood tests for
PT screening for the early stages of the disease.
XX
XX Claim 1; Page 160-161; 242pp; Japanese.
XX
XX The invention relates to stomach cancer-expressed genes (AAI93842-
CC AAI93917) and the encoded proteins (AAM93967-AAM94039). The genes can be
CC used as markers in blood tests for screening for the early stages of the
CC disease. The proteins and peptides can be used as targets for screening
CC for compounds to treat the disease. They can also be used for predicting
CC micro-metastases. The gene can predict peritoneal dissemination
XX
XX Sequence 589 AA;
XX
XX Query Match 44.3%; Score 51; DB 4; Length 589;
XX Best Local Similarity 60.0%; Pred. No. 59;
XX Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 2 TAPAAEGAGAEVKKASAEAK 21
DB 16 SAAGAEGAGAPAAAASAEPK 35
XX
XX RESULT 15
XX ADE59377
XX ID ADE59377 standard; protein; 589 AA.
XX AC ADE59377;
XX
XX 29-JAN-2004 (first entry)
XX
XX Human Protein NP_038466, SEQ ID NO 5271.
XX
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
XX Homo sapiens.
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEO) GEN HOSPITAL CORP.
PA (PARE) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
DR GENBANK; NP_038466.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
```

```
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 589 AA;
XX
XX Query Match 44.3%; Score 51; DB 7; Length 589;
XX Best Local Similarity 60.0%; Pred. No. 59;
XX Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 2 TAPAAEGAGAEVKKASAEAK 21
DB 16 SAAGAEGAGAPAAAASAEPK 35
XX
XX Search completed: October 7, 2004, 17:34:01
XX Job time : 55.5 secs
```

US-10-802-644-2
; Sequence 2, Application US/10802644

; Publication No. US20040180836A1

; GENERAL INFORMATION:
; APPLICANT: Martin, Linda
; APPLICANT: Li, Yuehua
; TITLE OF INVENTION: BLOCKING PEPTIDE FOR INFLAMMATORY CELL SECRETION
; FILE REFERENCE: 5051.574
; CURRENT APPLICATION NUMBER: US/10/802,644
; CURRENT FILING DATE: 2004-03-17
; PRIOR APPLICATION NUMBER: US/10/180,753
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: N-terminal myristoylation
US-10-802-644-2

Query Match 100.0%; Score 115; DB 16; Length 24;
Best Local Similarity 100.0%; Pred. No. 4.3e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTAPAAEGAGAEVVKRAAEAKQAF 24
|||||
Db 1 GTAPAAEGAGAEVVKRAAEAKQAF 24

RESULT 3

US-10-437-963-147454
; Sequence 147454, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 147454
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_47981C.1.1.pep
US-10-437-963-147454

Query Match 46.1%; Score 53; DB 16; Length 406;
Best Local Similarity 52.6%; Pred. No. 35;
Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTAPAAEGAGAEVVKRAAE 19
|||||
Db 6 GGTPEEGGGADVERAAAE 24

RESULT 4

US-10-282-122A-50399
; Sequence 50399, Application US/10282122A

; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA_034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50399
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Burkholderia mallei
US-10-282-122A-50399

Query Match 46.1%; Score 53; DB 12; Length 487;
Best Local Similarity 53.6%; Pred. No. 42;
Matches 15; Conservative 4; Mismatches 5; Indels 4; Gaps 2;

QY 1 GTA-PAAGAGA---EVKRAAEAKQAF 24
|||||
Db 24 GTGEPVWEGAGASADDERAVASARRAF 51

RESULT 5

US-10-437-963-178225
; Sequence 178225, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963

/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 178225
/ LENGTH: 116
/ TYPE: PRT
/ ORGANISM: Oryza sativa
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(116)
/ OTHER INFORMATION: unsure at all Xaa locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRI4530_75801C.1.pep
US-10-437-963-178225

Query Match 45.2%; Score 52; DB 16; Length 116;
Best Local Similarity 55.0%; Pred. No. 13;
Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 GTAPAAEGAGAEVKRASAEA 20
Db 96 GDRPAEGSGEAAALAA 115

RESULT 6
US-10-767-701-34508
/ Sequence 34508, Application US/10767701
/ Publication No. US20040172684A1
/ GENERAL INFORMATION:
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53535)B
/ CURRENT APPLICATION NUMBER: US/10767,701
/ CURRENT FILING DATE: 2004-01-29
/ NUMBER OF SEQ ID NOS: 63128
/ SEQ ID NO 34508
/ LENGTH: 177
/ TYPE: PRT
/ ORGANISM: Sorghum bicolor
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(177)
/ OTHER INFORMATION: unsure at all Xaa locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: SORBI-28MAY03-C4969_1.pep
US-10-767-701-34508

Query Match 45.2%; Score 52; DB 16; Length 177;
Best Local Similarity 52.4%; Pred. No. 20;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 APAAGAGAEVKRASAEAKQA 23
Db 101 SPATGAAAKQKQASIMAKQS 121

RESULT 7
US-10-282-122A-45228
/ Sequence 45228, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Yamamoto, Robert

/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 45228
/ LENGTH: 409
/ TYPE: PRT
/ ORGANISM: Acinetobacter baumannii
US-10-282-122A-45228

Query Match 44.3%; Score 51; DB 12; Length 409;
Best Local Similarity 60.0%; Pred. No. 67;
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 APAAGAGAEVKRASAEAKQ 22
Db 269 APSGEGATRCIKLALAEAKQ 288

RESULT 8
US-10-282-122A-75108
/ Sequence 75108, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06

```
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 75108
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Salmonella typhimurium
US-10-282-122A-75108

Query Match 44.3%; Score 51; DB 12; Length 488;
Best Local Similarity 54.5%; Pred. No. 80;
Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 3 APAAGGAGAEVKRASAEAKQAF 24
Db 31 AEVAGGAEVNVQVAAAEKAF 52

RESULT 9
US-10-293-000-4
; Sequence 4, Application US/10293000
; Publication No. US20030175278A1
; GENERAL INFORMATION:
; APPLICANT: Monteiro, Mervyn J.
; APPLICANT: Mah, Alex L.
; APPLICANT: Perry, George
; APPLICANT: Smith, Mark A.
; TITLE OF INVENTION: UBILQUILIN, A PRESENILIN INTERACTOR AND METHODS OF USING SAME
; FILE REFERENCE: 4115-175
; CURRENT APPLICATION NUMBER: US/10/293,000
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: US 60/338,549
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 4
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-000-4

Query Match 44.3%; Score 51; DB 14; Length 589;
Best Local Similarity 60.0%; Pred. No. 98;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 TAPAAEGAGAEVKRASAEAK 21
Db 16 SAAGAGAGAPAAASAEK 35

RESULT 10
US-10-293-000-2
; Sequence 2, Application US/10293000
; Publication No. US20030175278A1
; GENERAL INFORMATION:
; APPLICANT: Monteiro, Mervyn J.
; APPLICANT: Mah, Alex L.
; APPLICANT: Perry, George
; APPLICANT: Smith, Mark A.
; TITLE OF INVENTION: UBILQUILIN, A PRESENILIN INTERACTOR AND METHODS OF USING SAME
; FILE REFERENCE: 4115-175
```

```
; CURRENT APPLICATION NUMBER: US/10/293,000
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: US 60/338,549
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 2
; LENGTH: 595
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-000-2

Query Match 44.3%; Score 51; DB 14; Length 595;
Best Local Similarity 60.0%; Pred. No. 99;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 TAPAAEGAGAEVKRASAEAK 21
Db 16 SAAGAGAGAPAAASAEK 35

RESULT 11
US-10-156-761-14054
; Sequence 14054, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14054
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14054

Query Match 43.9%; Score 50.5; DB 14; Length 224;
Best Local Similarity 40.6%; Pred. No. 42;
Matches 13; Conservative 4; Mismatches 6; Indels 9; Gaps 1;

QY 1 GTAPAAEG-----AGAEVKRASAEAKQA 23
Db 124 GSAPAAAGTGAGSGSSAGSATTESAAKQA 155

RESULT 12
US-10-425-114-47199
; Sequence 47199, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
```


; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 47199
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3079-043-H2_FLI.pep
US-10-425-114-47199

Query Match 43.5%; Score 50; DB 12; Length 298;
Best Local Similarity 50.0%; Pred. No. 66;
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 3 APAEAGAGAEVVKASAEAKQ 22
Db 182 APAEAGAGAAVQGGGERR 201

RESULT 13
US-10-156-761-14429
; Sequence 14429, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14429
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14429

Query Match 43.5%; Score 50; DB 14; Length 395;
Best Local Similarity 60.0%; Pred. No. 89;
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 GTAPAAEGAGAEVVKASAEA 20
Db 116 GTGPAVEQAVALVERAAKEA 135

RESULT 14
US-09-864-408A-8324
; Sequence 8324, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shinketsu, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8324
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-864-408A-8324

Query Match 43.0%; Score 49.5; DB 11; Length 74;
Best Local Similarity 41.9%; Pred. No. 18;
Matches 13; Conservative 5; Mismatches 4; Indels 9; Gaps 1;

QY 2 TAPAAEAGA-----GAEVKASAEAKQA 23
Db 33 TAPVGAAGVAAVMASGTGALELVVAAAEAKA 63

RESULT 15
US-10-437-963-165448
; Sequence 165448, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 165448
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(143)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_64253C.1.pep
US-10-437-963-165448

Query Match 42.6%; Score 49; DB 16; Length 143;
Best Local Similarity 43.8%; Pred. No. 42;
Matches 14; Conservative 0; Mismatches 8; Indels 10; Gaps 1;

QY 1 GTAPAAEGAGAEVVKRA-----SAEAKQ 22
Db 93 GDRPAAEGGGVAATRAATAVAKRLGGSAAKQ 124

Search completed: October 7, 2004, 17:50:36
Job time : 86.5 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:32:03 ; Search time 15 Seconds
(without alignments)
82.602 Million cell updates/sec

Title: US-10-802-644-2

Perfect score: 115

Sequence: 1 GTAPAAEGAGAEVKEASAEAKQAF 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A-COMB.pep.*
- 2: /cgn2_6/ptodata/2/1aa/5B-COMB.pep.*
- 3: /cgn2_6/ptodata/2/1aa/6A-COMB.pep.*
- 4: /cgn2_6/ptodata/2/1aa/6B-COMB.pep.*
- 5: /cgn2_6/ptodata/2/1aa/PTUS-COMB.pep.*
- 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	45.2	211	4	US-09-252-991A-24706
2	52	45.2	407	4	US-09-252-991A-31200
3	52	45.2	453	4	US-09-252-991A-24290
4	51	44.3	422	4	US-09-328-352-4812
5	51	44.3	589	4	US-09-665-479A-4
6	50	43.5	259	4	US-09-252-991A-23293
7	50	43.5	492	4	US-09-489-039A-9023
8	49	42.6	173	4	US-09-252-991A-27742
9	49	42.6	418	4	US-09-328-352-5700
10	48.5	42.2	163	4	US-09-252-991A-21344
11	48.5	42.2	197	4	US-09-252-991A-22076
12	48	41.7	402	4	US-09-252-991A-29857
13	47.5	41.3	181	4	US-09-328-352-4739
14	47	40.9	31	4	US-09-690-454-159
15	47	40.9	125	4	US-09-690-454-155
16	47	40.9	143	1	US-08-253-155A-46
17	47	40.9	143	4	US-08-253-039A-12453
18	47	40.9	165	4	US-09-252-991A-17356
19	47	40.9	168	4	US-09-489-039A-7835
20	47	40.9	218	3	US-08-675-885-5
21	47	40.9	256	4	US-09-252-991A-26244
22	47	40.9	317	4	US-09-328-352-6383
23	47	40.9	329	4	US-09-252-991A-21954
24	47	40.9	378	3	US-09-032-372-11
25	47	40.9	378	3	US-08-853-733B-2
26	47	40.9	378	3	US-08-675-885-7
27	47	40.9	498	4	US-09-489-039A-7569

28 47 40.9 610 4 US-09-252-991A-20299 Sequence 20299, A
29 46 40.0 136 4 US-09-732-210-814 Sequence 814, App
30 46 40.0 136 4 US-09-732-210-815 Sequence 815, App
31 46 40.0 219 2 US-08-557-309B-54 Sequence 54, Appl
32 46 40.0 249 4 US-09-252-991A-30556 Sequence 30556, A
33 46 40.0 284 3 US-08-564-164A-2 Sequence 2, Appl
34 46 40.0 442 3 US-08-834-306-52 Sequence 52, Appl
35 46 40.0 442 3 US-08-993-674A-52 Sequence 52, Appl
36 46 40.0 442 4 US-09-256-976-52 Sequence 52, Appl
37 46 40.0 500 4 US-09-354-123-6 Sequence 1, Appl
38 46 40.0 571 2 US-08-796-414B-1 Sequence 1, Appl
39 46 40.0 584 4 US-09-198-452A-352 Sequence 352, App
40 46 40.0 612 4 US-09-252-991A-25727 Sequence 25727, A
41 46 40.0 767 4 US-09-252-991A-28484 Sequence 28484, A
42 45.5 39.6 388 4 US-09-252-991A-30608 Sequence 30608, A
43 45 39.1 84 4 US-09-489-039A-10806 Sequence 10806, A
44 45 39.1 444 3 US-09-252-292C-27 Sequence 27, Appl
45 45 39.1 472 4 US-09-252-991A-29165 Sequence 29165, A

ALIGNMENTS

RESULT 1
US-09-252-991A-24706
; Sequence 24706, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24706
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24706

Query Match 45.2%; Score 52; DB 4; Length 211;
Best Local Similarity 52.2%; Pred. No. 3.7;
Matches 12; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 GTAPAAEGAGAEVKEASAEAKQA 23
Db 79 GSAPATRRAGAAADRAAARAAA 101

RESULT 2
US-09-252-991A-31200
; Sequence 31200, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31200
; LENGTH: 407
; TYPE: PRT

```
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31200

Query Match      45.2%; Score 52; DB 4; Length 407;
Best Local Similarity 59.1%; Pred. No. 7.5;
Matches 13; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY      2 TAPAAEGAGAEVVKRAAEAKQ 23
      :|||:|||||:|||||:
Db      220 SATAAEGGAGAEQARAAREAGAA 241

RESULT 3
US-09-252-991A-24290
; Sequence 24290, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24290
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24290

Query Match      45.2%; Score 52; DB 4; Length 453;
Best Local Similarity 64.7%; Pred. No. 8.5;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      3 APAAEGAGAEVVKRAAE 19
      :|||:|||||:|||||:
Db      421 APAAEGGAGAEQQAQAAD 437

RESULT 4
US-09-328-352-4812
; Sequence 4812, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4812
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4812

Query Match      44.3%; Score 51; DB 4; Length 422;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      3 APAAEGAGAEVVKRAAEAKQ 22
      :|||:|||||:|||||:
Db      282 APSGEGATRCIKLALAEAKQ 301

RESULT 5
US-09-665-479A-4
; Sequence 4, Application US/09665479A
; Patent No. 6673570
; GENERAL INFORMATION:
; APPLICANT: Itoh, Fumiko
; APPLICANT: Itoh, Susumu
; APPLICANT: Heldin, Carl-Henrik
; APPLICANT: ten-Dijke, Peter
; TITLE OF INVENTION: SMAD ASSOCIATING POLYPEPTIDES
; FILE REFERENCE: L00461.70096.US
; CURRENT APPLICATION NUMBER: US/09/665,479A
; CURRENT FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US 60/154,846
; PRIOR FILING DATE: 1999-09-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-665-479A-4

Query Match      44.3%; Score 51; DB 4; Length 589;
Best Local Similarity 60.0%; Pred. No. 16;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY      2 TAPAAEGAGAEVVKRAAEAK 21
      :|||:|||||:|||||:
Db      16 SAAGAEGAGAPAAAAAAREPK 35

RESULT 6
US-09-252-991A-23293
; Sequence 23293, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23293
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23293

Query Match      43.5%; Score 50; DB 4; Length 259;
Best Local Similarity 47.8%; Pred. No. 9;
Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY      1 GTPAAEGAGAEVVKRAAEAKQ 23
      :|||:|||||:|||||:
Db      152 GNAPAGGAGAEERSRIARHLEA 174

RESULT 7
US-09-489-039A-9023
; Sequence 9023, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
```

```

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A

```

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29857
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29857

Query Match 41.7%; Score 48; DB 4; Length 402;
Best Local Similarity 47.8%; Pred. No. 28;
Matches 11; Conservative 3; Mismatches 9; Indels 9; Gaps 0;

Qy 1 GTAPAAEGAGAEVKRASAERKQA 23
Db 26 GGDPSVAGGAGVQRAPGEAVRA 48

RESULT 13
US-09-328-352-4739
; Sequence 4739, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTG99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4739
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4739

Query Match 41.3%; Score 47.5; DB 4; Length 181;
Best Local Similarity 48.0%; Pred. No. 14;
Matches 12; Conservative 3; Mismatches 5; Indels 5; Gaps 1;

Qy 5 AAEAGAGAEVKRAS-----AEAKQAF 24
Db 152 AAEAGAGAEVKRAS-----AEAKQAF 176

RESULT 14
US-09-690-454-159
; Sequence 159, Application US/09690454
; Patent No. 6531447
; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al.
; TITLE OF INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: PZ006P1
; CURRENT APPLICATION NUMBER: US/09/690,454
; CURRENT FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: 09/189,144
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 60/044,039
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,093
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,190
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/050,935
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,101
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,356
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/056,250
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,296
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,293
; PRIOR FILING DATE: August 29, 1997
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 155
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (30)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-690-454-155

Query Match 40.9%; Score 47; DB 4; Length 125;
Best Local Similarity 76.9%; Pred. No. 11;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 AEGAGAEVKRASA 18
Db 6 AEGAGAEVETATA 18

RESULT 15
US-09-690-454-155
; Sequence 155, Application US/09690454
; Patent No. 6531447
; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al.
; TITLE OF INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: PZ006P1
; CURRENT APPLICATION NUMBER: US/09/690,454
; CURRENT FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: 09/189,144
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 60/044,039
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,093
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,190
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/050,935
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,101
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,356
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/056,250
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,296
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,293
; PRIOR FILING DATE: August 29, 1997
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 155
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (30)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-690-454-155

Query Match 40.9%; Score 47; DB 4; Length 125;
Best Local Similarity 76.9%; Pred. No. 11;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 AEGAGAEVKRASA 18
Db 6 AEGAGAEVETATA 18

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29857
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29857

Query Match 41.7%; Score 48; DB 4; Length 402;
Best Local Similarity 47.8%; Pred. No. 28;
Matches 11; Conservative 3; Mismatches 9; Indels 9; Gaps 0;

Qy 1 GTAPAAEGAGAEVKRASAERKQA 23
Db 26 GGDPSVAGGAGVQRAPGEAVRA 48

RESULT 13
US-09-328-352-4739
; Sequence 4739, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTG99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4739
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4739

Query Match 41.3%; Score 47.5; DB 4; Length 181;
Best Local Similarity 48.0%; Pred. No. 14;
Matches 12; Conservative 3; Mismatches 5; Indels 5; Gaps 1;

Qy 5 AAEAGAGAEVKRAS-----AEAKQAF 24
Db 152 AAEAGAGAEVKRAS-----AEAKQAF 176

RESULT 14
US-09-690-454-159
; Sequence 159, Application US/09690454
; Patent No. 6531447
; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al.
; TITLE OF INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: PZ006P1
; CURRENT APPLICATION NUMBER: US/09/690,454
; CURRENT FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: 09/189,144
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 60/044,039
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,093
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,190
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/050,935
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,101
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,356
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/056,250
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,296
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,293
; PRIOR FILING DATE: August 29, 1997
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 155
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (30)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-690-454-155

Query Match 40.9%; Score 47; DB 4; Length 125;
Best Local Similarity 76.9%; Pred. No. 11;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 AEGAGAEVKRASA 18
Db 6 AEGAGAEVETATA 18

RESULT 15
US-09-690-454-155
; Sequence 155, Application US/09690454
; Patent No. 6531447
; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al.
; TITLE OF INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: PZ006P1
; CURRENT APPLICATION NUMBER: US/09/690,454
; CURRENT FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: 09/189,144
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 60/044,039
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,093
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,190
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/050,935
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,101
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,356
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/056,250
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,296
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,293
; PRIOR FILING DATE: August 29, 1997
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 155
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (30)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-690-454-155

Query Match 40.9%; Score 47; DB 4; Length 125;
Best Local Similarity 76.9%; Pred. No. 11;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 AEGAGAEVKRASA 18
Db 6 AEGAGAEVETATA 18

Tue Oct 12 05:37:43 2004

us-10-802-644-2.open.ra1

Page 5

Db 43 AEGAGAEVETATA 5S

Search completed: October 7, 2004, 17:36:36
Job time : 17 secs

This Page Blank (uspto)


```

Query Match      23.5%; Score 27; DB 2; Length 22;
Best Local Similarity 29.4%; Pred. No. 1.9e+03;
Matches 5; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY      5 SKTAAKGFAAAARPGEA 21
      : : : : :
DB      4 AKSGGNEVGDDADPGDA 20

```

```
RESULT 3
PS0209
24K protein 4407 - rice (strain Nihonbare) (fragment)
C;Species: Oryza sativa (rice)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 11-Apr-1995
C;Accession: PS0209
R;Tsugita, A.; Kamo, M.
Submitted to JIPID, April 1993
A;Reference number: PS0209
A;Accession: PS0209
A;Molecule type: protein
A;Residues: 1-10 <TSU>
A;Experimental source: callus
C;Comment: molecular weight 24K, pI 4.6.

Query Match      22.6%; Score 26; DB 2; Length 10;
Best Local Similarity 75.0%; Pred. No. 1.3e+03;
Matches      6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      15 AERPGEAA 22
      ||| |||
Db      3 AEVPAEAA 10

RESULT 4
A49177
22K protein p1, mitochondrial - Crithidia fasciculata (fragment)
C;Species: Crithidia fasciculata
C;Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 19-May-1994
C;Accession: A49177; PH1826
R;Tittawella, I.
Exp. Cell Res. 206, 143-151, 1993
A;Title: Identification of DNA-binding proteins in the parasitic protozoan Crithidia fasciculata
A;Reference number: A49177; PMID:93245912; PMID:8482355
A;Accession: A49177
A;Molecule type: protein
A;Residues: 1-15 <TIT>
A;Note: sequence extracted from NCBI backbone (NCBIP:1311148)
C;Comment: This protein binds with mitochondrial DNA.
C;Keywords: DNA binding; mitochondrion

Query Match      22.6%; Score 26; DB 2; Length 15;
Best Local Similarity 46.2%; Pred. No. 1.9e+03;
Matches      6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy      12 EAARERPGEEAAV 24
      :||| |||
Db      1 DAPASAPKAAAA 13

RESULT 5
PQ0688
Photosystem I 14.0K E4 chain - common tobacco (fragment)
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 17-Mar-1999
C;Accession: PQ0688
R;Obokata, J.; Mikami, K.; Hayashida, N.; Nakamura, M.; Sugiyura, M.
Plant Physiol. 102, 1259-1267, 1993
A;Title: Molecular heterogeneity of photosystem I. psad, psae, psaf, psah and psal are a
A;Reference number: PQ0667; PMID:94105345; PMID:8278548
A;Accession: PQ0688
A;Molecule type: protein
A;Residues: 1-20 <OBO>
C;Keywords: chloroplast; photosynthesis; photosystem I

Query Match      22.6%; Score 26; DB 2; Length 20;
Best Local Similarity 40.0%; Pred. No. 2.4e+03;
Matches      8; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

Qy      5 SKTAAGGAARPGEEAAV 24
      :||| |||
Db      1 AEEAAPPAATATAGEAPPA 20
```

RESULT 6

```
T25653
hypothetical protein C47C12.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T25653
R;Connell, M.
Submitted to the EMBL Data Library, August 1996
A;Description: The sequence of C. elegans cosmid C47C12.
A;Reference number: Z20062
A;Accession: T25653
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-22 <CON>
A;Cross-references: EMBL:U67951; PIDN:AAB07573.1; GSPDB:GNC00028; CESP:C47C12.5
A;Experimental source: strain Bristol N2; clone C47C12
C;Genetics:
A;Gene: CESP:C47C12.5
A;Map position: X
A;Introns: 21/1
```

```
Query Match      22.6%; Score 26; DB 2; Length 22;
Best Local Similarity 50.0%; Pred. No. 2.6e+03;
Matches      5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

Qy 3 QFSKTAAKGE 12

Db 12 QFSRISLKN 21

RESULT 7

```
S68900
xanthine dehydrogenase (EC 1.1.1.204) 82.4K chain - Veillonella atypica (fragment)
C;Species: Veillonella atypica
C;Date: 23-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 29-Aug-1997
C;Accession: S68900
R;Grewer, L.; Meyer, O.
Eur. J. Biochem. 238, 862-866, 1996
A;Title: Characterization of xanthine dehydrogenase from the anaerobic bacterium Veillonella
A;Reference number: S68900; PMID:96300255; PMID:8706691
A;Accession: S68900
A;Molecule type: protein
A;Residues: 1-22 <GRE>
A;Experimental source: DSM 1399
C;Keywords: 2Fe-2S; FAD; flavoprotein; heterotrimer; iron-sulfur protein; metalloprotein
```

```
Query Match      22.6%; Score 26; DB 2; Length 22;
Best Local Similarity 45.5%; Pred. No. 2.6e+03;
Matches      5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

Qy 1 GAQFSKTAAGK 11

Db 5 GKSVDKVDAGK 15

RESULT 8

```
B48186
ATP synthase beta-2 chain - wood tobacco (fragment)
C;Species: Nicotiana sylvestris (wood tobacco)
C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 23-Mar-1995
C;Accession: B48186
R;De Paeppe, R.; Forchioni, A.; Chetrit, P.; Vedel, F.
Proc. Natl. Acad. Sci. U.S.A. 90, 5934-5938, 1993
A;Title: Specific mitochondrial proteins in pollen: presence of an additional ATP synthase
A;Reference number: A48186; PMID:93317598; PMID:8327463
A;Accession: B48186
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-23 <DEI>
A;Experimental source: pollen
A;Note: sequence extracted from NCBI backbone (NCBIP:134868)
```

Query Match 22.6%; Score 26; DB 2; Length 23;
 Best Local Similarity 83.3%; Pred. No. 2.7e+03;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 13 AAAERP 18
 |||||
 Db 4 AAAERP 9

RESULT 9
 C30221
 histone H2A.8 - chicken (fragment)
 C:Species: Gallus gallus (Chicken)
 C>Date: 02-Nov-1989 #sequence_revision 02-Nov-1989 #text_change 30-Sep-1993
 C:Accession: C30221
 R:Challoner, P.B.; Moss, S.B.; Groudine, M.
 Mol. Cell. Biol. 9, 902-913, 1989
 A>Title: Expression of replication-dependent histone genes in avian spermatids involves
 A:Reference number: A30221; MUID:89261754; PMID:2471062
 A:Accession: C30221
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-17 <CHA>

Query Match 21.7%; Score 25; DB 2; Length 17;
 Best Local Similarity 46.2%; Pred. No. 2.9e+03;
 Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGEA 13
 |||||
 Db 5 GKQGGKARAKA 17

RESULT 10
 A41299
 T-cell receptor alpha chain precursor V region (17.2) - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 23-Jul-1999
 C:Accession: A41299
 R:Uematsu, Y.; Wege, H.; Straus, A.; Ott, M.; Bannwarth, W.; Lanchbury, J.; Panayi, G.;
 Proc. Natl. Acad. Sci. U.S.A. 88, 8534-8538, 1991
 A>Title: The T-cell-receptor repertoire in the synovial fluid of a patient with rheumatoid
 A:Reference number: A41299; MUID:92020887; PMID:1656449
 A:Accession: A41299
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-19 <UEM>
 A:Cross-references: GB:S57433; MID:g236318; PIDN:AA819956.1; PID:g236319
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: T-cell receptor

Query Match 21.7%; Score 25; DB 2; Length 19;
 Best Local Similarity 57.1%; Pred. No. 3.2e+03;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 17 RPGEAAV 23
 |||||
 Db 9 QPGDSAV 15

RESULT 11
 PQ0687
 photosystem I 14.1K E3 chain - common tobacco (fragment)
 C:Species: Nicotiana tabacum (common tobacco)
 C>Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 17-Mar-1999
 C:Accession: PQ0687; PQ0675
 R:Obokata, J.; Mikami, K.; Hayashida, N.; Nakamura, M.; Sugiyura, M.
 Plant Physiol. 102, 1259-1267, 1993
 A>Title: Molecular heterogeneity of photosystem I. psaD, psaE, psaF, psaH and psaL are
 A:Reference number: PQ0687; MUID:94105345; PMID:8278548
 A:Accession: PQ0687
 A:Molecule type: protein
 A:Residues: 1-20 <OBO>

A>Note: this peptide was designated photosystem I 14.1K E3 chain
 A:Accession: PQ0675
 A:Molecule type: protein
 A:Residues: 2-11 <OB2>
 A>Note: this peptide was designated photosystem I 14.0K E4 chain
 C:Keywords: chloroplast; photosynthesis; photosystem I; thylakoid

Query Match 21.7%; Score 25; DB 2; Length 20;
 Best Local Similarity 41.2%; Pred. No. 3.3e+03;
 Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 5 SKTAAGGEAAERFGEA 21
 :|||
 Db 2 AEEAAPPAATATAGEA 18

RESULT 12
 S65611
 tubulin beta chain - bovine (fragment)
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 07-May-1999
 C:Accession: S65611
 R:Ruediger, M.; Plessmann, U.; Ruediger, A.H.; Weber, K.
 FEBS Lett. 364, 147-151, 1995
 A>Title: Beta tubulin of bull sperm is polyglycylated.
 A:Reference number: S65611; MUID:95269788; PMID:7750559
 A:Accession: S65611
 A:Molecule type: protein
 A:Residues: 1-21 <RUE>
 C:Complex: heterodimer; alpha and beta chain
 C:Superfamily: tubulin
 C:Keywords: GTP binding; heterodimer; microtubule

Query Match 21.7%; Score 25; DB 2; Length 21;
 Best Local Similarity 55.6%; Pred. No. 3.5e+03;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 12 EAAAEERPE 20
 :|||
 Db 12 DATAEEGE 20

RESULT 13
 PC7043
 ubiquitin carboxyl-terminal hydrolase (EC 3.1.1.-) homolog - yeast (Pichia anomala) (frag
 N:Alternate names: ALX1 protein
 C:Species: Pichia anomala, Candida pelliculosa
 C>Date: 04-Mar-2000 #sequence_revision 04-Mar-2000 #text_change 11-May-2000
 C:Accession: PC7043
 R:Sakajo, S.; Minagawa, N.; Yoshimoto, A.
 Biosci. Biotechnol. Biochem. 63, 1889-1894, 1999
 A>Title: Structure and regulatory expression of a single copy alternative oxidase gene fr
 A:Reference number: JC7145; MUID:20101471; PMID:10635554
 A:Accession: PC7043
 A:Molecule type: DNA
 A:Residues: 1-21 <SAK>
 A:Cross-references: DDBJ:AB026726
 C:Genetics:
 A:Gene: alx1
 C:Keywords: hydrolase

Query Match 21.7%; Score 25; DB 2; Length 21;
 Best Local Similarity 37.5%; Pred. No. 3.5e+03;
 Matches 6; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 5 SKTAAGGEAAERFGEA 20
 :|||
 Db 3 AKEAKVGSVEDKEGE 18

RESULT 14
 H30608
 Ig kappa chain V-III region (Ste) - human (fragment)

C;Species: Homo sapiens (man)
C;Date: 29-Jun-1989 #sequence_revision 29-Jun-1989 #text_change 30-May-1997
C;Accession: H30608
R;Goni, F.R.; Chen, P.P.; McGinnis, D.; Arjonilla, M.L.; Fernandez, J.; Carson, D.; Soldo
J. Immunol. 142, 3158-3163, 1989
A;Title: Structural and idiotypic characterization of the L chains of human IGM autoantibodies
A;Reference number: A30601; MUID:89215279; PMID:2496160

A;Accession: H30608
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-22 <GON>
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

Query Match 21.7%; Score 25; DB 2; Length 22;
Best Local Similarity 71.4%; Pred. No. 3.6e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 18 PGEAAVA 24
||| |
Db 15 PGERATA 21

RESULT 15

PQ0143
polygalacturonase (EC 3.2.1.15) P26 - evening primrose (fragment)
C;Species: Oenothera organensis (evening primrose)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 06-Dec-1996
C;Accession: PQ0143

R;Brown, S.M.; Crouch, M.L.
Plant Cell 2, 263-274, 1990
A;Title: Characterization of a gene family abundantly expressed in Oenothera organensis
A;Reference number: JQ0992; MUID:93005658; PMID:2152116

A;Accession: PQ0143
A;Molecule type: mRNA
A;Residues: 1-22 <BRO>
A;Experimental source: pollen
C;Comment: This protein is specifically translated in the pollens.
C;Comment: This protein functions by depolymerizing pectin in the cell walls of the pistil
Y growing tube.
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 21.7%; Score 25; DB 2; Length 22;
Best Local Similarity 50.0%; Pred. No. 3.6e+03;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 5 SKTAAKGEAA 14
:| ||| |
Db 12 TKYAKGDGA 21

Search completed: October 7, 2004, 17:57:24
Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:36:42 ; Search time 14 Seconds
(without alignments)
89.263 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GAQFSKTAAGKGAARPGCAAVA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 1550

Minimum DB seq length: 0

Maximum DB seq length: 24

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	27.0	21	1 MISG MISAN	P81474 misgurnus a
2	27	23.5	15	1 UC23 MAIZE	P80629 zea mays (m
3	27	23.5	22	1 YQX8 CAEEL	Q09290 caenorhabdi
4	26	22.6	18	1 FIBA LAMGL	P14454 lama glama
5	25	21.7	24	1 BOWN BOMVA	P01505 bombina var
6	24	20.9	14	1 MCR2 METTM	P58816 methanobact
7	24	20.9	16	1 FIBA TAPTE	P14536 capirus ter
8	24	20.9	17	1 FIBA PIG	P14460 sus scrofa
9	24	20.9	23	1 IBP2 PIG	P24853 sus scrofa
10	23	20.0	16	1 DBH3 RHILE	P80605 rhizobium l
11	23	20.0	19	1 FIBA BUBBU	P14442 bubalus bub
12	23	20.0	20	1 COXQ THUOB	P80983 thunnus obe
13	23	20.0	20	1 PORC METTM	P80902 methanobact
14	23	20.0	24	1 FK84 RAT	Q9qvc8 rattus norv
15	22	19.1	18	1 FIBA CAMDR	P14444 camelus dro
16	22	19.1	20	1 CRP MUSCA	P19094 mustelus ca
17	22	19.1	23	1 VG22 BPT2	P21596 bacterioph
18	22	19.1	23	1 VG22 BPT6	P21597 bacterioph
19	22	19.1	24	1 LPER STRFR	P45440 streptomyce
20	21.5	18.7	19	1 ANP7 ELEGR	P11920 eleginus gr
21	21.5	18.7	20	1 MCRG METTE	P22950 methanosarc
22	21	18.3	8	1 RS1 ERWCH	P37985 erwinia chr
23	21	18.3	10	1 COXQ THUOB	P80982 thunnus obe
24	21	18.3	12	1 H2AX ONCMY	P83327 oncorhynch
25	21	18.3	13	1 PSBP PINPS	P81668 pinus pinas
26	21	18.3	14	1 IF2G RAT	P81795 rattus norv
27	21	18.3	16	1 FIBA ODOHE	P14459 odocoileus
28	21	18.3	17	1 NEF HVLJ3	P12480 human immun
29	21	18.3	17	1 RM35 YEAST	P36530 saccharomyc
30	21	18.3	19	1 PHLC STAIN	P80924 staphylococ
31	21	18.3	19	1 UP24 UPEIN	P82030 uperoleia i
32	21	18.3	20	1 CD4 SHEEP	P05542 oviss aries
33	21	18.3	20	1 COXQ ONCMY	P80335 oncorhynch

34 20 17.4 9 1 DSIP RABIT
35 20 17.4 11 1 Q2OA COMTE
36 20 17.4 11 1 TKNA SCYCA
37 20 17.4 13 1 PSAB PEA
38 20 17.4 14 1 UC15 MAIZE
39 20 17.4 14 1 UHA1 CANFA
40 20 17.4 15 1 ATP2 SPIOL
41 20 17.4 16 1 ANP8 ELEGR
42 20 17.4 16 1 FIBA FELCA
43 20 17.4 16 1 FIBA HYLLA
44 20 17.4 19 1 FIBA BISBO
45 20 17.4 19 1 PHSL DESBN

ALIGNMENTS

RESULT 1
MISG MISAN STANDARD; PRT; 21 AA.
AC P81474;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Misgurn.
OS Misgurnus anguillicaudatus (Oriental weatherfish) (Oriental weatherloach).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cobitidae; Misgurnus.
OX NCBI_TaxID=75329;
RN [1]_TaxID=75329;
RP SEQUENCE.
RX MEDLINE=97415401; PubMed=9271200;
RA Park C.B., Lee H.J., Park I.Y., Kim M.S., Kim S.C.;
RT "A novel antimicrobial peptide from the loach, Misgurnus anguillicaudatus";
RL FEBS Lett. 411:173-178(1997).
CC -!- FUNCTION: Strong antimicrobial activity against several Gram-positive and Gram-negative bacteria and fungi.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MASS SPECTROMETRY: MW=2502; METHOD=MALDI.
KW Antibiotic; Fungicide.
SQ SEQUENCE 21 AA; 2502 MW; 4A6E9D0AB391BCF1 CRC64;
Query Match 27.0%; Score 31; DB 1; Length 21;
Best Local Similarity 46.7%; Pred. No. 2.3e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 3 QFSKTAAGKGAAR 17
Db 6 ELSKFSKGAARRR 20
RESULT 2
UC23 MAIZE STANDARD; PRT; 15 AA.
AC P80629;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Unknown protein from 2D-page of etiolated coleoptile (Spot 502) (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]_TaxID=4577;
RP SEQUENCE.
RC TISSUE=Coleoptile;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C., Pernollet J.-C., Zivy M., de Vienne D.;

```

RT "The maize two dimensional gel protein database: towards an integrated
RL genome analysis program.";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC protein is: 6.2, its MW is: 45.0 kDa.
DR Maize-2DPAGE; P80629; COLLEOPTILE.
DR MaizeDB; 123955; -.
FT NON_TER 1 1
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1557 MW; C974ED33E9A4EC28 CRC64;

Query Match 23.5%; Score 27; DB 1; Length 15;
Best Local Similarity 50.0%; Pred. No. 6.3e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 14 AAERPGEA 21
| :|:|:|
Db 1 AGDRPGDA 8

RESULT 3
YQK8_CABEL STANDARD; PRT; 22 AA.
AC Q03290;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 2.2 kDa protein C56G2.8 in chromosome III.
GN C56G2.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Connell M.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U23177; AAA64335.1; -.
DR PIR; T15870; T15870.
DR WormPep; C56G2.8; CE01877.
KW Hypothetical protein.
SQ SEQUENCE 22 AA; 2164 MW; 7585F9684FE19D51 CRC64;

Query Match 23.5%; Score 27; DB 1; Length 22;
Best Local Similarity 29.4%; Pred. No. 8.9e+02;
Matches 5; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 5 SKTAAGGEAAERPGEA 21
| :|:|:|
Db 4 AKSGGENVEGDADPGDA 20

RESULT 4
FIBA_LAMGL STANDARD; PRT; 18 AA.
AC F14454;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).
GN FGA.
OS Lama glama (Llama), and
OS Lama vicugna (Vicugna) (Vicugna vicugna).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.
OX NCBI_TaxID=9844, 9843;
RN [1]
RP SEQUENCE.
RC SPECIES=L.glama;
RA Blomback B., Blomback M., Grondahl N.J.;
RT "Studies on fibrinopeptides from mammals.";
EL Acta Chem. Scand. 19:1789-1791(1965).
RN [2]
RP SEQUENCE.
RC SPECIES=L.vicugna;
RA Mross G.A., Doolittle R.F.;
RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";
EL Arch. Biochem. Biophys. 122:674-684(1967).
CC -!- FUNCTION: Fibrinogen has a double function: yielding monomers that
CC polymerize into fibrin and acting as a cofactor in platelet
CC aggregation.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
CC which cleaves fibrinopeptides A and B from alpha and beta chains,
CC and thus exposes the N-terminal polymerization sites responsible
CC for the formation of the soft clot.
KW Blood coagulation; Plasma.
FT PEPTIDE 1 18 FIBRINOPEPTIDE A.
FT NON_TER 18 18
SQ SEQUENCE 18 AA; 1834 MW; 2444487B8B7F4CC6 CRC64;

Query Match 22.6%; Score 26; DB 1; Length 18;
Best Local Similarity 58.3%; Pred. No. 1e+03;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 8 AAKGEAAERPGE 19
| :|:|:|
Db 5 ADGGEFLAEGGG 16

RESULT 5
BOMN_BOMVA STANDARD; PRT; 24 AA.
ID BOMN_BOMVA
AC P01505;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Bombinin.
OS Bombina variegata (Yellow-bellied toad).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Archeobatrachia; Bombinatoridae; Bombina.
OX NCBI_TaxID=8348;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RA Csordas A., Michl H.;
RT "Isolation and structural resolution of a haemolytically active
RT polypeptide from the immune secretion of a European toad.";
EL Monatsh. Chem. 101:182-189(1970).
CC -!- FUNCTION: Has antimicrobial and hemolytic activities.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
CC -!- SIMILARITY: Belongs to the bombinin family.
DR PIR; A01766; BMTD.
KW Amphibian defense peptide; Antibiotic; Hemolysis; Amidation.
FT MOD_RES 24 24 AMIDATION.
SQ SEQUENCE 24 AA; 2294 MW; ACC0FCE3B7402E85 CRC64;

Query Match 21.7%; Score 25; DB 1; Length 24;
Best Local Similarity 53.8%; Pred. No. 1.9e+03;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GAQFSKTAAGGEA 13
| :|:|:|
Db 3 GALSAGKALGLA 15

```

```

RESULT 6
MCRZ_METTMM
ID MCRZ_METTMM STANDARD; PRT; 14 AA.
AC P58816;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Methyl-coenzyme M reductase II gamma subunit (EC 1.8.-.-) (MCR II
DE Gamma) (Fragment).
GN MRTG.
OS Methanobacterium thermoautotrophicum (strain Marburg / DSM 2133).
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=79929;
RN [1]
RP SEQUENCE.
RX MEDLINE=91099370; PubMed=2269306;
RA Rospert S., Linder D., Ellermann J., Thauer R.K.;
RT "Two genetically distinct methyl-coenzyme M reductases in
RT Methanobacterium thermoautotrophicum strain Marburg and delta H.";
RL Eur. J. Biochem. 194:871-877(1990).
CC -!- FUNCTION: Reduction of methyl-coenzyme M (2-(methylthio)
CC ethanesulfonic acid) with 7-mercaptoheptanoylthreonine phosphate
CC to methane and an heterodisulfide.
CC -!- CATALYTIC ACTIVITY: CH(3)-S-CoM + H-S-HTP = CH(4) + CoM-S-S-HTP.
CC -!- COFACTOR: Binds 2 coenzyme F430 noncovalently per hexamer.
CC Coenzyme F430 is a yellow nickel porphyrinoid.
CC -!- PATHWAY: Methanogenesis; last step.
CC -!- SUBUNIT: Hexamer of two alpha, two beta, and two gamma chains.
CC -!- DEVELOPMENTAL STAGE: THERE ARE TWO MCR COMPLEXES IN THIS BACTERIA.
CC MCR II IS EXPRESSED IN THE EARLY GROWTH PHASE. LATE GROWTH CELLS
CC CONTAINS MOSTLY MCR I.
DR PIR; E69017; E69017.
KW Methanogenesis; Oxidoreductase; Multigene family.
FT INIT MET 0 0
FT NON TER 14 14
SQ SEQUENCE 14 AA; 1557 MW; 97E9439C4223B871 CRC64;

Query Match 20.9%; Score 24; DB 1; Length 14;
Best Local Similarity 57.1%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 18 PGEAAVA 24
DB 8 PGETQIA 14

RESULT 7
FIBA_TAPTE
ID FIBA_TAPTE STANDARD; PRT; 16 AA.
AC P14536;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).
GN FGA.
OS Tapirus terrestris (Lowland tapir) (Brazilian tapir).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Tapiridae; Tapirus.
OX NCBI_TaxID=9801;
RN [1]
RP SEQUENCE.
RA O'Neil P.B., Doolittle R.F.;
RT "Mammalian phylogeny based on fibrinopeptide amino acid sequences.";
RL Syst. Zool. 22:590-595(1973).
CC -!- FUNCTION: Fibrinogen has a double function: yielding monomers that
CC polymerize into fibrin and acting as a cofactor in platelet
CC aggregation.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
CC which cleaves fibrinopeptides A and B from alpha and beta chains,
CC and thus exposes the N-terminal polymerization sites responsible
CC for the formation of the soft clot.
DR PIR; E29501; E29501.
KW Blood coagulation; Plasma.
FT PEPTIDE 1 16
FT NON TER 17 17
SQ SEQUENCE 17 AA; 1762 MW; 232EFE8B8B6B0A0C CRC64;

Query Match 20.9%; Score 24; DB 1; Length 17;
Best Local Similarity 60.0%; Pred. No. 1.9e+03;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 10 KGEAAAEPRG 19
DB 6 KGEFLAEGGG 15

RESULT 8
FIBA_PIG
ID FIBA_PIG STANDARD; PRT; 17 AA.
AC P14460;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).
GN FGA.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE.
RA Blomback B., Blomback M., Grondahl N.J.;
RT "Studies on fibrinopeptides from mammals.";
RL Acta Chem. Scand. 19:1789-1791(1965).
CC -!- FUNCTION: Fibrinogen has a double function: yielding monomers that
CC polymerize into fibrin and acting as a cofactor in platelet
CC aggregation.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
CC which cleaves fibrinopeptides A and B from alpha and beta chains,
CC and thus exposes the N-terminal polymerization sites responsible
CC for the formation of the soft clot.
DR PIR; E29501; E29501.
KW Blood coagulation; Plasma.
FT PEPTIDE 1 16
FT NON TER 17 17
SQ SEQUENCE 17 AA; 1762 MW; 232EFE8B8B6B0A0C CRC64;

Query Match 20.9%; Score 24; DB 1; Length 17;
Best Local Similarity 60.0%; Pred. No. 1.9e+03;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 10 KGEAAAEPRG 19
DB 6 KGEFLAEGGG 15

RESULT 9
IBP2_PIG
ID IBP2_PIG STANDARD; PRT; 23 AA.
AC P24853;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Insulin-like growth factor binding protein 2 (IGFBP-2) (IBP-2)
DE (IGF-binding protein 2) (Fragment).
GN IGFBP2.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;

```

CC which cleaves fibrinopeptides A and B from alpha and beta chains,
 CC and thus exposes the N-terminal polymerization sites responsible
 CC for the formation of the soft clot.

KW Blood coagulation; Plasma.
 FT PEPTIDE 1 16
 FT NON TER 16 16
 SQ SEQUENCE 16 AA; 1622 MW; 48598EB6292F4030 CRC64;

Query Match 20.9%; Score 24; DB 1; Length 16;
 Best Local Similarity 50.0%; Pred. No. 1.8e+03;
 Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 8 AAKGEAAAEPRG 19
 DB 3 ATEGEFLAEGGG 14

RESULT 8
 FIBA_PIG
 ID FIBA_PIG STANDARD; PRT; 17 AA.
 AC P14460;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).
 GN FGA.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE.
 RA Blomback B., Blomback M., Grondahl N.J.;
 RT "Studies on fibrinopeptides from mammals.";
 RL Acta Chem. Scand. 19:1789-1791(1965).

CC -!- FUNCTION: Fibrinogen has a double function: yielding monomers that
 CC polymerize into fibrin and acting as a cofactor in platelet
 CC aggregation.
 CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -!- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
 CC which cleaves fibrinopeptides A and B from alpha and beta chains,
 CC and thus exposes the N-terminal polymerization sites responsible
 CC for the formation of the soft clot.

DR PIR; E29501; E29501.
 KW Blood coagulation; Plasma.
 FT PEPTIDE 1 16
 FT NON TER 17 17
 SQ SEQUENCE 17 AA; 1762 MW; 232EFE8B8B6B0A0C CRC64;

Query Match 20.9%; Score 24; DB 1; Length 17;
 Best Local Similarity 60.0%; Pred. No. 1.9e+03;
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 10 KGEAAAEPRG 19
 DB 6 KGEFLAEGGG 15

RESULT 9
 IBP2_PIG
 ID IBP2_PIG STANDARD; PRT; 23 AA.
 AC P24853;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Insulin-like growth factor binding protein 2 (IGFBP-2) (IBP-2)
 DE (IGF-binding protein 2) (Fragment).
 GN IGFBP2.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;

```

RN  [1]
RP  SEQUENCE.
RX  MEDLINE=92109718; PubMed=1722398;
RA  Coleman M.E., Pan Y.-C.E., Etherton T.D.;
RT  "Identification and NH2-terminal amino acid sequence of three
RT  insulin-like growth factor-binding proteins in porcine serum.";
RL  Biochem. Biophys. Res. Commun. 181:1131-1136(1991).
CC  -!- FUNCTION: IGF-binding proteins prolong the half-life of the IGFs
CC  and have been shown to either inhibit or stimulate the growth
CC  promoting effects of the IGFs on cell culture. They alter the
CC  interaction of IGFs with their cell surface receptors.
CC  -!- SUBCELLULAR LOCATION: Secreted.
CC  -!- MISCELLANEOUS: Binds IGF-II more than IGF-I.
CC  -!- SIMILARITY: Contains 1 IGFBP domain.
CC  -!- SIMILARITY: Contains 1 thyroglobulin type-I domain.
DR  PIR; JH0515; JH0515.
DR  InterPro; IPR000967; Insl_gro_fac_pr.
DR  InterPro; IPR000716; Thyroglobulin 1.
DR  PROSITE; PS00222; IGF BINDING; PARTIAL.
DR  PROSITE; PS00484; THYROGLOBULIN_1; PARTIAL.
KW  Growth factor binding.
FT  NON TER 23
SQ  SEQUENCE 23 AA; 2387 MW; 731104B88825D2C4 CRC64;

Query Match 20.9%; Score 24; DB 1; Length 23;
Best Local Similarity 66.7%; Pred. No. 2.5e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 14 AAERPGEAA 22
DB 12 AAXRPPPA 20
||| |||
||| |||

RESULT 10
DBH3 RHILE
ID DBH3 RHILE STANDARD; PRT; 16 AA.
AC P80605;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA-binding protein H3-RL (Fragment).
OS Rhizobium leguminosarum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=384;
RN [1]
RP SEQUENCE.
RC STRAIN=CWR 538;
RX MEDLINE=20035754; PubMed=10570977;
RA Sojda J. III, Gu B., Lee J., Hoover T.R., Nixon B.T.;
RT "A rhizobial homolog of IHF stimulates transcription of dctA in
RT Rhizobium leguminosarum but not in Sinorhizobium meliloti.";
RL Gene 238:489-500(1999).
CC -!- FUNCTION: Histone-like DNA-binding protein which is capable of
CC wrapping DNA to stabilize it, and thus to prevent its denaturation
CC under extreme environmental conditions.
CC -!- SIMILARITY: Belongs to the bacterial histone-like protein family.
DR InterPro; IPR000119; Bac_DNAbind
DR Pfam; PF00216; Bac_DNA_binding_1.
DR PROSITE; PS00045; HISTONE LIKE; PARTIAL.
KW DNA-binding; DNA condensation.
FT NON TER 16
SQ SEQUENCE 16 AA; 1702 MW; F585ADED46B33964 CRC64;

Query Match 20.0%; Score 23; DB 1; Length 16;
Best Local Similarity 71.4%; Pred. No. 2.5e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 13 AAAERPG 19
DB 9 AVAERAG 15
||| |||
||| |||

us-10-802-644-1.closed.rsp
```

```

RESULT 11
FIBA_BUBBU STANDARD; PRT; 19 AA.
AC P14442; P14443;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).
GN FGA.
OS Bubalus bubalis (Domestic water buffalo).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bubalus.
OX NCBI_TaxID=89462;
RN [1]
RP SEQUENCE.
RA Moss G.A., Doolittle R.F.;
RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";
RL Arch. Biochem. Biophys. 122:674-684(1967).
RN [2]
RP SEQUENCE.
RC STRAIN=Italian breed;
RX MEDLINE=76040091; PubMed=1180969;
RA Balestrieri C., Colonna G., Irace G.;
RT "Covalent structure of fibrinopeptides from buffaloes breeding in
RT Italy.";
RL Biochim. Biophys. Acta 405:517-521(1975).
CC -!- FUNCTION: Fibrinogen has a double function: Yielding monomers that
CC polymerize into fibrin and acting as a cofactor in platelet
CC aggregation.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
CC which cleaves fibrinopeptides A and B from alpha and beta chains,
CC and thus exposes the N-terminal polymerization sites responsible
CC for the formation of the soft clot.
KW Blood coagulation; Plasma.
FT PEPTIDE 1 19 FIBRINOPEPTIDE A.
FT VARIANT 8 8 S -> G (IN ITALIAN BRED).
FT NON TER 19
SQ SEQUENCE 19 AA; 1852 MW; 9BA41F0F55A54CC5 CRC64;

Query Match 20.0%; Score 23; DB 1; Length 19;
Best Local Similarity 50.0%; Pred. No. 2.9e+03;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 8 AAKCEAAARPDG 19
DB 6 AVSGEFLAEGG 17
||| |||
||| |||

RESULT 12
COXQ_THUOB STANDARD; PRT; 20 AA.
ID COXQ_THUOB
AC P80983;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome c oxidase polypeptide VIII-1 (EC 1.9.3.1) (Fragment).
OS Thunnus obesus (Bigeye tuna).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC Scombridae; Thunnus.
OX NCBI_TaxID=8241;
RN [1]
RP SEQUENCE.
RC TISSUE=Heart, and Liver;
RX MEDLINE=97454291; PubMed=9310366;
RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
RA Kadenbach B.;
RT "The subunit structure of cytochrome-c oxidase from tuna heart and
RT liver.";
```



```

RL Eur. J. Biochem. 248:99-103(1997).
CC -!- FUNCTION: This protein is one of the nuclear-coded polypeptide
CC chains of cytochrome c oxidase, the terminal oxidase in
CC mitochondrial electron transport.
CC -!- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -!- SIMILARITY: Belongs to the cytochrome c oxidase VIII family.
DR PIR; S77991; S77991.
KW Oxidoreductase; Mitochondrion.
FT NON TER 20
SQ SEQUENCE 20 AA; 1993 MW; 5BFEA5080A862C58 CRC64;

Query Match 20.0%; Score 23; DB 1; Length 20;
Best Local Similarity 35.3%; Pred. No. 3e+03;
Matches 6; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 8 AAKGEAAARPCGAFAVA 24
DB 2 SAKPAKXXVTAGEQAIA 18

RESULT 13
PQRC_METTM STANDARD; PRT; 20 AA.
AC P80502;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Pyruvate synthase subunit PQRC (EC 1.2.7.1) (Pyruvate oxidoreductase
DE gamma chain) (POR) (Pyruvic-ferredoxin oxidoreductase gamma subunit)
DE (Fragment).
GN PORC.
OS Methanobacterium thermoautotrophicum (strain Marburg / DSM 2133).
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=79929;
RN [1]
RP SEQUENCE.
RX MEDLINE=97261844; PubMed=9108258;
RA Tersteegen A., Linder D., Thauer R.K., Hedderich R.;
RT "Structures and functions of four anabolic 2-oxoacid oxidoreductases
RT in Methanobacterium thermoautotrophicum.";
RL Eur. J. Biochem. 244:862-868(1997).
CC -!- CATALYTIC ACTIVITY: Pyruvate + CoA + oxidized ferredoxin = acetyl-
CC CoA + CO(2) + reduced ferredoxin.
CC -!- COFACTOR: Binds 2 4Fe-4S clusters (By similarity).
CC -!- SUBUNIT: Heterotetramer of one alpha, one beta, one delta and one
CC gamma chain.
CC -!- MISCELLANEOUS: Has a pH optimum of 10.0 and an optimal temperature
CC of 80 degrees Celsius.
CC -!- SIMILARITY: Belongs to the bacterial-type ferredoxin family.
DR InterPro; IPR002869; POR.
KW Oxidoreductase; Electron transport; Iron-sulfur; Repeat; 4Fe-4S.
FT NON TER 20
SQ SEQUENCE 20 AA; 2167 MW; 481532134D2F353 CRC64;

Query Match 20.0%; Score 23; DB 1; Length 20;
Best Local Similarity 50.0%; Pred. No. 3e+03;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAQFSKTAAK 10
DB 10 GGQXAVTAAE 19

RESULT 14
FKB4_RAT STANDARD; PRT; 24 AA.
AC Q9QVC8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE FK506-binding protein 4 (EC 5.2.1.8) (Peptidyl-prolyl cis-trans

```

```

DE isomerase) (PPIase) (Rotamase) (p59 protein) (HSP binding
DE immunophilin) (HBI) (FKBP52 protein) (52 kDa FK506 binding protein)
DE (FKBP59) (Fragment).
GN FKBp4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RC TISSUE=Leukemia;
RX MEDLINE=92285692; PubMed=1376003;
RA Tai P.-K.K., Albers M.W., Chang H., Faber L.E., Schreiber S.L.;
RT "Association of a 59-kilodalton immunophilin with the glucocorticoid
RT receptor complex.";
RL Science 256:1315-1318(1992).
CC -!- FUNCTION: Component of unactivated mammalian steroid receptor
CC complexes that sediment at 8-10 S. May have a rotamase activity.
CC May play a role in the intracellular trafficking of hetero-
CC oligomeric forms of steroid hormone receptors (By similarity).
CC -!- CATALYTIC ACTIVITY: Peptidylproline (omega=180) = peptidylproline
CC (omega=0).
CC -!- SUBUNIT: Associates with HSP90 and HSP70 in unactivated steroid
CC hormone receptor complexes. Also interacts with peroxisomal
CC phytanoyl-CoA alpha-hydroxylase (PHYH). Interacts with NR3C1 and
CC dynein (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the FKBP-type PPIase family.
DR InterPro; IPR001179; FKBP_PPIase.
DR PROSITE; PS00453; FKBP_PPIASE_1; PARTIAL.
DR PROSITE; PS00454; FKBP_PPIASE_2; PARTIAL.
DR PROSITE; PS00059; FKBP_PPIASE_3; PARTIAL.
KW Isomerase; Rotamase; Nuclear protein.
FT UNSURE 10
FT UNSURE 14
FT UNSURE 18
FT UNSURE 20
FT NON TER 24
SQ SEQUENCE 24 AA; 2478 MW; 901A4B67B2AF5D16 CRC64;

Query Match 20.0%; Score 23; DB 1; Length 24;
Best Local Similarity 35.3%; Pred. No. 3.6e+03;
Matches 6; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 2 AQFSKTAAGGAARERP 18
DB 2 AEEMKVAENGAQSAPLP 18

RESULT 15
FKB4_CAMDR STANDARD; PRT; 18 AA.
AC F14444;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).
GN FGA.
OS Camelus dromedarius (Dromedary) (Arabian camel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.
OX NCBI_TaxID=9838;
RN [1]
RP SEQUENCE.
RX MEDLINE=67209145; PubMed=6033721;
RA Doolittle R.F., Schubert D., Schwartz S.A.;
RT "Amino acid sequence studies on artiodactyl fibrinopeptides. I.
RT Dromedary camel, mule deer, and cape buffalo.";
RL Arch. Biochem. Biophys. 118:456-467(1967).
CC -!- FUNCTION: Fibrinogen has a double function: yielding monomers that
CC polymerize into fibrin and acting as a cofactor in platelet
CC aggregation.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS

```

CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -1- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
 CC which cleaves fibrinopeptides A and B from alpha and beta chains,
 CC and thus exposes the N-terminal polymerization sites responsible
 CC for the formation of the soft clot.
 KW Blood coagulation; Plasma.
 FT PEPTIDE 1 18 FIBRINOPEPTIDE A.
 FT NON TER 18 18
 SQ SEQUENCE 18 AA; 1835 MW; 244448763D7F4CC6 CRC64;
 Query Match 19.1%; Score 22; DB 1; Length 18;
 Best Local Similarity 50.0%; Pred. No. 3.8e+03;
 Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 8 AAKGEMAAERPG 19
 Db 5 ADEGEFLAEGGG 16

Search completed: October 7, 2004, 17:54:02
 Job time : 17 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:37:47 ; Search time 72 Seconds

(without alignments)
105.173 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GAQFSKTAAGAAARPGERA VA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 10229

Minimum DB seq length: 0

Maximum DB seq length: 24

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	25.2	15	10 Q9SQI6	Q9sqi6 oryza sativ
2	29	25.2	20	11 Q99JV2	Q99jv2 mus musculu
3	29	25.2	21	10 Q9S8K6	Q9s8k6 mesembryant
4	29	25.2	23	4 Q9BYM6	Q9bym6 homo sapien
5	28	24.3	17	3 Q00074	Q00074 aspergillus
6	28	24.3	20	2 Q51558	Q51558 pseudomonas
7	28	24.3	20	10 Q9SM31	Q9sm31 triticum ae
8	28	24.3	22	8 Q9T217	Q9t217 nicotiana s
9	27	23.5	21	4 Q86UF6	Q86uf6 homo sapien
10	27	23.5	22	11 Q8CJD5	Q8cjd5 rattus norv
11	27	23.5	22	12 Q85664	Q85664 reovirus (t
12	26	22.6	13	6 Q9TQ52	Q9tq52 equus cabal
13	26	22.6	15	5 Q9TW04	Q9tw04 crithidia f
14	26	22.6	18	11 Q55167	Q55167 rattus norv
15	26	22.6	20	2 Q85522	Q85522 chlamydia t
16	26	22.6	20	2 Q85523	Q85523 chlamydia t

17	26	22.6	20	2 Q85528	Q85528 chlamydia t
18	26	22.6	20	2 Q85512	Q85512 chlamydia t
19	26	22.6	20	2 Q85531	Q85531 chlamydia t
20	26	22.6	20	2 Q85510	Q85510 chlamydia t
21	26	22.6	20	2 Q85521	Q85521 chlamydia t
22	26	22.6	20	2 Q85514	Q85514 chlamydia t
23	26	22.6	20	2 Q85519	Q85519 chlamydia t
24	26	22.6	20	2 Q85517	Q85517 chlamydia t
25	26	22.6	20	2 Q85525	Q85525 chlamydia t
26	26	22.6	20	2 Q85526	Q85526 chlamydia t
27	26	22.6	20	2 Q69176	Q69176 chlamydia t
28	26	22.6	20	2 Q85520	Q85520 chlamydia t
29	26	22.6	20	2 Q85533	Q85533 chlamydia t
30	26	22.6	20	2 Q85509	Q85509 chlamydia t
31	26	22.6	20	2 Q85530	Q85530 chlamydia t
32	26	22.6	20	8 Q9T219	Q9t219 nicotiana s
33	26	22.6	20	15 Q78508	Q78508 human immun
34	26	22.6	21	6 Q9TR75	Q9tr75 bos taurus
35	26	22.6	22	5 Q94182	Q94182 caenorhabdi
36	26	22.6	22	6 Q9TRB7	Q9trb7 bos taurus
37	26	22.6	23	2 Q84F40	Q84f40 streptomyce
38	26	22.6	23	8 Q9T2S5	Q9t2s5 nicotiana s
39	25	21.7	8	5 P82688	P82688 periplaneta
40	25	21.7	13	2 Q9REI2	Q9rei2 acidiphilia
41	25	21.7	16	3 Q9UR86	Q9ur86 candida par
42	25	21.7	19	6 Q9TR32	Q9tr32 ovis aries
43	25	21.7	19	13 Q91433	Q91433 gallus gall
44	25	21.7	20	8 Q9T2I8	Q9t2i8 nicotiana s
45	25	21.7	20	15 Q78505	Q78505 human immun

ALIGNMENTS

RESULT 1

Q9SQI6 PRELIMINARY; PRT; 15 AA.
ID AC Q9SQI6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Translation elongation factor (Fragment).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Li Z.-Y., Chen S.-Y.;
RT "Inducible expression of translation elongation factor 1A gene in rice seedlings in response to environmental stresses.";
RL Acta Bot. Sin. 41:800-806(1999).
DR EMBL; AF067195; AAC79991.1; -.
DR Gramene; Q9SQI6; -.
DR GO; GO:0003746; F:translation elongation factor activity; IEA.
KW Elongation factor.
FT NON TER 1 1
SQ SEQUENCE 15 AA; 1514 MW; 95E5C46F069C9775 CRC64;

Query Match 25.2%; Score 29; DB 10; Length 15;
Best Local Similarity 60.0%; Pred. No. 2e+03;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAQFSKTAAK 10
||: ||||
Db 4 GAKVTKAAAX 13

RESULT 2

Q99JV2 PRELIMINARY; PRT; 20 AA.
ID AC Q99JV2

DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC005653; AA005653.1; -;
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 20 AA; 2223 MW; F203F6A80A7A0429 CRC64;

 Query Match 25.2%; Score 29; DB 11; Length 20;
 Best Local Similarity 62.5%; Pred. No. 2.7e+03;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

 QY 13 AAAERPG 20
 Db :|||:|
 3 SPAKPGK 10

 RESULT 3
 Q9S8K6 PRELIMINARY; PRT; 21 AA.
 ID Q9S8K6;
 AC Q9S8K6;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE 24 kDa salt-responsive glycoprotein/group 5 pathogenesis-related
 DE protein homolog (Fragment).
 OS Mesembryanthemum crystallinum (Common ice plant).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllales; Alzooaceae; Mesembryanthemum.
 OX NCBI_TaxID=3544;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=95062715; PubMed=7972493;
 RA Yen H.E., Edwards G.E., Grimes H.D.;
 RT "Characterization of a salt-responsive 24-kilodalton glycoprotein in
 RT Mesembryanthemum crystallinum.";
 RL Plant Physiol. 105:1179-1187(1994).
 SQ SEQUENCE 21 AA; 2288 MW; B463C490647E076D CRC64;

 Query Match 25.2%; Score 29; DB 10; Length 21;
 Best Local Similarity 85.7%; Pred. No. 2.9e+03;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 QY 13 AAAERPG 19
 Db :|||:|
 15 AAAARPG 21

 RESULT 4
 Q9BYM6 PRELIMINARY; PRT; 23 AA.
 ID Q9BYM6;
 AC Q9BYM6;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE DSRNA adenosine deaminase (Fragment).
 GN DSRNA
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;
 RA Deblandre G., Marinx O., Nols C., Defrance P., Berr P., Huez G.,
 RA Caput D.;
 RT "The gene coding for the interferon-inducible human dsRNA adenosine
 RT deaminase is transcribed into several messengers specifying different
 RT proteins.";
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X98560; CAA67171.1; -;
 FT NON_TER 23
 SQ SEQUENCE 23 AA; 2543 MW; CF29179B7DFC1395 CRC64;

 Query Match 25.2%; Score 29; DB 4; Length 23;
 Best Local Similarity 50.0%; Pred. No. 3.1e+03;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

 QY 10 KGEAAARPG 19
 Db :|||:|
 5 RNESAAGKPG 14

 RESULT 5
 Q00074 PRELIMINARY; PRT; 17 AA.
 ID Q00074;
 AC Q00074;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Extracellular acid protease (Fragment).
 GN PEPA.
 OS Aspergillus niger.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5061;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=N400;
 RX MEDLINE=95163121; PubMed=7532112;
 RA Jarai G., Buxton F.;
 RT "Nitrogen, carbon, and pH regulation of extracellular acidic proteases
 RT of Aspergillus niger.";
 RL Curr. Genet. 26:238-244 (1994).
 DR EMBL; U03507; AAC48920.1; -;
 DR MEROPS; A01.016; -;
 DR GO; GO:0008233; F:peptidase activity; IEA.
 KW Protease.
 FT NON_TER 17
 SQ SEQUENCE 17 AA; 1694 MW; FF01D0EBC9FEAE0B CRC64;

 Query Match 24.3%; Score 28; DB 3; Length 17;
 Best Local Similarity 100.0%; Pred. No. 3.2e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 4 FSKTAA 9
 Db :|||:|
 4 FSKTAA 9

 RESULT 6
 Q51558 PRELIMINARY; PRT; 20 AA.
 ID Q51558;
 AC Q51558;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Protein 10 (77).
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90130290; PubMed=2105298;
 RA Hancock A.N., Iglewski B.H.;

RT "Expression of the Pseudomonas aeruginosa toxA positive regulatory
 RT gene (regA) in Escherichia coli.";
 RL 7. Bacteriol. 172:589-594 (1990).
 DR EMBL; M33819; AAA25982.1; -.
 SQ SEQUENCE 20 AA; 2158 MW; 8878365B519BFE4A CRC64;

Query Match 24.3%; Score 28; DB 2; Length 20;
 Best Local Similarity 37.5%; Pred. No. 3.8e+03;
 Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 5 SNTAAKGEAAARPGGE 20
 ::|||:|
 Db 2 ARIRAGSSRACGLRGE 17

RESULT 7

Q9SM31 PRELIMINARY; PRT; 20 AA.
 AC Q9SM31;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE Putative sucrose synthase (Fragment).
 GN SS1.
 OS Triticum aestivum (Wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Triticeae; Triticum.
 OX NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Chinese spring;
 RA Helguera M., Khan I.A., Dubcovsky J.;
 RT "Development of PCR markers for wheat leaf rust resistance gene
 Lr47.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ249623; CAB61261.1; -.
 FT NON TER 1
 SQ SEQUENCE 20 AA; 2298 MW; CC4EFE0F3877CCDC CRC64;

Query Match 24.3%; Score 28; DB 10; Length 20;
 Best Local Similarity 83.3%; Pred. No. 3.8e+03;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 ERPGEA 21
 |:|||
 Db 15 EQPGEA 20

RESULT 8

Q9T217 PRELIMINARY; PRT; 22 AA.
 AC Q9T217;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE 14.3 photosystem I PSAE protein (Fragment).
 OS Nicotiana sylvestris (Wood tobacco).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Lamiales; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4096;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=94105345; PubMed=8278548;
 RA Obokata J., Mikami K., Hayaashida N., Nakamura M., Sugiura M.;
 RT "Molecular heterogeneity of photosystem I. psaD, psaE, psaF, psaH, and
 psaL are all present in isoforms in Nicotiana spp.";
 RL Plant Physiol. 102:1259-1267 (1993).
 SQ SEQUENCE 22 AA; 2033 MW; 488E17845223FFF7 CRC64;

Query Match 24.3%; Score 28; DB 8; Length 22;

Best Local Similarity 38.9%; Pred. No. 4.2e+03;
 Matches 7; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
 QY 6 KTAAXGEAAARPGGAAY 23
 :|:|||
 Db 3 EAAPPAAATAEPAPAPV 20

RESULT 9

Q86UF6 PRELIMINARY; PRT; 21 AA.
 AC Q86UF6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Betacellulin (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nakagawa T.;
 RT "The 5' end of human betacellulin gene and molecular scanning for
 RT mutations in Japanese patients with type2 diabetes mellitus.";
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB109442; BAC76610.1; -.
 FT NON TER 21
 SQ SEQUENCE 21 AA; 2083 MW; 8387517EAA8B9FEB CRC64;

Query Match 23.5%; Score 27; DB 4; Length 21;
 Best Local Similarity 41.2%; Pred. No. 5.5e+03;
 Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 8 AAKGEAAARPGGAAY 24
 ||:|||
 Db 4 AARGGASSLPALLALA 20

RESULT 10

Q8CJD5 PRELIMINARY; PRT; 22 AA.
 AC Q8CJD5;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Podocin (Fragment).
 GN NPHS2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Morita H., Yoshimura A., Makino R., Inui K., Nakao N., Usami T.,
 RA Roselli S., Antignac C., Matsuyama M., Ideura T.;
 RT "Rat genome fragment containing a part of exons and all of the 3'UTR
 of Nphs2 as well as microsatellite sites.";
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB094123; BAC23093.1; -.
 FT NON TER 1
 SQ SEQUENCE 22 AA; 2313 MW; DBC687F825956D93 CRC64;

Query Match 23.5%; Score 27; DB 11; Length 22;
 Best Local Similarity 75.0%; Pred. No. 5.8e+03;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 AAKGEAAA 15
 ||:|||
 Db 11 AAEGEKAA 18

```

RESULT 11
Q85664
ID Q85664 PRELIMINARY; PRT; 22 AA.
AC Q85664;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Major core protein lambda-1 (Fragment).
OS Reovirus (type 3 / strain Dearing).
OC Viruses; dsRNA viruses; Reoviridae; Orthoreovirus.
OX NCBI_TaxID=10886;
RN [1]
RP SEQUENCE OF 1-6 FROM N.A.
RC STRAIN=Dearing;
RX MEDLINE=82217029; PubMed=7086967;
RA Kozak M.;
RT "Sequences of ribosome binding sites from the large size class of reovirus mRNA.";
RL J. Virol. 42:467-473 (1982).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Dearing;
RX MEDLINE=83017876; PubMed=6927854;
RA Antczak J.B., Chmelo R., Pickup D.J., Joklik W.K.;
RT "Sequences at both termini of the 10 genes of reovirus serotype 3 (strain Dearing).";
RL Virology 121:307-319 (1982).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Dearing;
RX MEDLINE=83017877; PubMed=7123853;
RA Gaillard R.K., Li J.K., Keene J.D., Joklik W.K.;
RT "The sequences at the termini of four genes of the three reovirus serotypes.";
RL Virology 121:320-326 (1982).
DR EMBL; J02317; AAA47254.1; -.
FT NON_TER 22
SQ SEQUENCE 22 AA; 2462 MW; D6706A861D930E1F CRC64;

Query Match 23.5%; Score 27; DB 12; Length 22;
Best Local Similarity 41.7%; Pred. No. 5.8e+03;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 KTAAGGEAAER 17
Db 11 KSSGKNDSTER 22

RESULT 12
Q9TQS2
ID Q9TQS2 PRELIMINARY; PRT; 13 AA.
AC Q9TQS2;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE Transferrin (Fragment).
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RC Giffard J.M., Brandon R.B., Bell T.K.;
RT "Further identification of single nucleotide polymorphisms in the equine transferrin gene.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF185752; AAF05466.1; -.
DR EMBL; AF185748; AAF05462.1; -.
DR EMBL; AF185749; AAF05463.1; -.
DR EMBL; AF185750; AAF05464.1; -.
DR EMBL; AF185751; AAF05465.1; -.
FT NON_TER 13
SQ SEQUENCE 13 AA; 1361 MW; D893D92670B6CEBD CRC64;

Query Match 22.6%; Score 26; DB 6; Length 13;
Best Local Similarity 71.4%; Pred. No. 4.7e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 17 RPGEAAV 23
Db 2 RFGSACV 8

RESULT 13
Q9TWU4
ID Q9TWU4 PRELIMINARY; PRT; 15 AA.
AC Q9TWU4;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE KDNA-binding protein (Fragment).
OS Crithidia fasciculata.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Crithidia.
OX NCBI_TaxID=5656;
RN [1]
RP SEQUENCE.
RX MEDLINE=93245912; PubMed=8482355;
RA Tittawella I.;
RT "Identification of DNA-binding proteins in the parasitic protozoan Crithidia fasciculata and evidence for their association with the mitochondrial genome.";
RL Exp. Cell Res. 206:143-151 (1993).
DR PIR; A49177; A49177.
SQ SEQUENCE 15 AA; 1453 MW; 7C737E529D9D8270 CRC64;

Query Match 22.6%; Score 26; DB 5; Length 15;
Best Local Similarity 46.2%; Pred. No. 5.5e+03;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 12 EAAAEPPGEAAVA 24
Db 1 DAPASAPKAAA 13

RESULT 14
O55167
ID O55167 PRELIMINARY; PRT; 18 AA.
AC O55167;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE ARE1 (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEW 1W/Gun;
RX MEDLINE=99009314; PubMed=9790748;
RA Walter L., Guenther E.;
RT "Identification of a novel highly conserved gene in the centromeric part of the major histocompatibility complex.";
RL Genomics 52:298-304 (1998).
DR EMBL; AJ223831; CAA11568.1; -.
FT NON_TER 18
SQ SEQUENCE 18 AA; 1505 MW; B88156A2048C3388 CRC64;

Query Match 22.6%; Score 26; DB 11; Length 18;
Best Local Similarity 53.3%; Pred. No. 6.6e+03;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 8 AAKGEAAERPGEEA 22
Db 11 KSSGKNDSTER 22

```

Db 4 AATMAAAAGAGVAA 18

RESULT 15

O85522 PRELIMINARY; PRT; 20 AA.
 AC O85522;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE Major outer membrane protein (Fragment).
 GN omp-1.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B23a;
 RA Hsieh Y.-H., Bobo L.D.;
 RT "Diversity of major outer membrane protein (omp-1) of Chlamydia trachomatis in trachoma endemic villages, Kongwa, Tanzania.";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF070275; AAC25245.1; -
 FT NON_TER 1
 FT NON_TER 20
 SQ SEQUENCE 20 AA; 1929 MW; 504C8985119AF1F3 CRC64;

Query Match 22.6%; Score 26; DR 2; Length 20;
 Best Local Similarity 66.7%; Pred. No. 7.3e+03;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 7 TAAKGEAAA 15
 |||||
 Db 4 TATTGNAAA 12

Search completed: October 7, 2004, 17:56:33
 Job time : 76 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:36:12 ; Search time 84.5 Seconds
(without alignments)
80.250 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GAQFSKTRAAKGAAERPGAAVA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 538975

Minimum DB seq length: 0

Maximum DB seq length: 24

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	100.0	24	3	AAY95896 MANS pept
2	115	100.0	24	6	ABP97765 Myristoyl
3	39	33.9	24	4	ABB40149 Peptide #
4	39	33.9	24	4	AAM33792 Peptide #
5	39	33.9	24	4	AAM73597 Human bon
6	39	33.9	24	4	AAM60912 Human bon
7	39	33.9	24	4	ABG55330 Human liv
8	39	33.9	24	5	ABG43467 Human pep
9	37	32.2	16	2	AAY30222 KEA16, am
10	37	32.2	16	2	AAY32964 Membrane
11	37	32.2	16	7	AAY38860 Membrane
12	37	32.2	19	2	AAY20179 Human bet
13	37	32.2	24	7	ADB81153 Human pan
14	36	31.3	20	2	AAY36444 Fragment
15	36	31.3	20	6	ADAL1979 Human nov
16	35	30.4	21	6	ABU06998 Maize Sta
17	35	30.4	21	6	ABU06998 Maize Sta
18	34	29.6	16	2	AAR04447 Human imm
19	34	29.6	16	2	AAY30220 AXE16, am
20	34	29.6	16	2	AAY32962 Membrane
21	34	29.6	16	7	AAY38858 Membrane
22	34	29.6	21	2	AAM19062 Trypanoso
23	34	29.6	21	2	AAM19087 Trypanoso
24	34	29.6	21	2	AAY32839 TCE repea
25	34	29.6	21	2	AAY23311 Epitope o

ALIGNMENTS

RESULT 1

AAY95896

ID AAY95896 standard; peptide; 24 AA.

XX

AC AAY95896;

XX

DT 20-NOV-2000 (first entry)

XX

DE MANS peptide inhibitor of MARCKS-related mucus secretion.

XX

MANS; MARCKS; myristoylated alanine-rich C kinase substrate; human;

mucus secretion; inhibitor; bronchitis; asthma; cystic fibrosis;

KW chronic obstructive pulmonary disease; pneumonia; emphysema; influenza;

KW rhinitis; therapy.

XX

OS Homo sapiens.

XX

PN WO2000050062-A2.

XX

PD 31-AUG-2000.

XX

PF 24-FEB-2000; 2000WO-US005050.

XX

PR 24-FEB-1999; 99US-00256154.

XX

PA (UYNC-) UNIV NORTH CAROLINA STATE.

XX

PI Li Y, Martin LD, Adler KB;

XX

WPI; 2000-572036/53.

XX

Regulating mucus secretion by a mucus-secreting cell, useful for treating

e.g. bronchitis, asthma or pneumonia, by administering a compound that

inhibits or enhances myristoylated alanine-rich C-kinase substrate

PT

protein.

XX

PS Claim 12; Page 40; 66pp; English.

XX

CC The present sequence is that of MANS peptide, comprising the N-terminal

region of human myristoylated alanine-rich C kinase substrate MARCKS

protein (see AAY95896), a major cellular substrate for protein kinase S.

CC MANS peptide inhibits secretion of mucus from mucus membranes and mucus-

secreting cells, including human airway epithelial cells. It is suggested

CC to block attachment of MARCKS protein to the mucin granule, thus blocking

CC or inhibiting the release of mucin granules and the secretion of mucus by

CC the cell. The invention relates to methods and compounds for decreasing

CC mucus secretion, particularly in the airways. Such compounds include MANS

peptide and antisense oligonucleotides to MARCKS. They are useful in

26 34 29.6 21 2 AAY23313 Repeat se
27 34 29.6 21 3 AAB26466 T.cruzi T
28 34 29.6 21 3 AAB26468 T.cruzi s
29 34 29.6 23 4 AAU03178 Partial p
30 33 28.7 15 3 AAY58971
31 33 28.7 15 3 AAY82060 MHC class
32 33 28.7 15 4 AAG63202 Peptide w
33 33 28.7 16 2 AAW30841 MyokL pro
34 33 28.7 16 2 AAW30221 EXA16, am
35 33 28.7 16 2 AAW30223
36 33 28.7 16 2 AAY32963 Membrane
37 33 28.7 16 2 AAY32965
38 33 28.7 16 5 ABG92842 Synthetic
39 33 28.7 16 5 ABG92827 Synthetic
40 33 28.7 16 5 ABG92508 Self-asse
41 33 28.7 16 5 ABG92493 Self-asse
42 33 28.7 16 6 ABP59360 Self-asse
43 33 28.7 16 6 ABP59376 Self-asse
44 33 28.7 16 7 AAE38859 Membrane
45 33 28.7 16 7 AAE38861 Membrane

CC inhibiting mucus secretion in conditions such as bronchitis, cystic
CC fibrosis, chronic obstructive pulmonary disease, asthma, emphysema,
CC pneumonia, influenza, rhinitis and the common cold
XX
SQ Sequence 24 AA;

Query Match 100.0%; Score 115; DB 3; Length 24;
Best Local Similarity 100.0%; Pred. No. 9.6e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAARPGGAAVA 24
|||
DB 1 GAQFSKTAAGGAAARPGGAAVA 24
|||

RESULT 2

ABP97765
ID ABP97765 standard; peptide; 24 AA.

XX AC

XX ABP97765;

XX DT 11-AUG-2003 (first entry)

XX XX

DE Myristoylated N-terminal sequence (MANS) peptide.

XX XX

KW Myristoylated N-terminal sequence; MANS; MARCKS; mucus secretion;

KW inflammatory mediator; inflammation; respiratory disease; asthma;

KW chronic bronchitis; chronic obstructive pulmonary disease; COPD;

KW bowel disease; irritable bowel syndrome; Crohn's disease;

KW ulcerative colitis; skin disease; rosacea; eczema; psoriasis; acne;

KW autoimmune disease; pain; arthritis; cystic fibrosis.

XX OS Synthetic.

XX XX

FH Key Location/Qualifiers

FT Modified-site 1

FT /note= "N-terminal myristate chain attached"

XX XX

PN WO2003000027-A2.

XX XX

PD 03-JAN-2003.

XX XX

PF 26-JUN-2002; 2002WO-US022270.

XX XX

PR 26-JUN-2001; 2001US-0300933P.

XX XX

PA (UUNC-) UNIV NORTH CAROLINA STATE.

XX XX

PI Martin LD, Adler KB, Li Y;

XX XX

DR WPI; 2003-278239/27.

XX XX

PT Method of regulating inflammation comprises administering a composition comprising a MANS peptide or an active fragment thereof.

XX XX

PS Disclosure; Page 26; 54pp; English.

XX XX

CC The present sequence represents a myristoylated N-terminal sequence (MANS) peptide. The MANS peptide is identical to the first 24 amino acids of MARCKS, and mediates insertion of MARCKS into membranes. The MANS peptide inhibits both mucus secretion and inflammatory mediators. The specification describes a method of regulating inflammation. The method comprises administering a composition comprising a MANS peptide. The method is useful for treating inflammation caused by respiratory diseases (e.g. asthma, chronic bronchitis and chronic obstructive pulmonary disease (COPD), bowel diseases (e.g. irritable bowel syndrome, Crohn's disease and ulcerative colitis), skin diseases (e.g. rosacea, eczema, psoriasis and severe acne), autoimmune diseases and pain syndromes, arthritis and cystic fibrosis

XX XX

SQ Sequence 24 AA;

Query Match 100.0%; Score 115; DB 6; Length 24;

Best Local Similarity 100.0%; Pred. No. 9.6e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAARPGGAAVA 24
|||
DB 1 GAQFSKTAAGGAAARPGGAAVA 24
|||

RESULT 3

ABB40149
ID ABB40149 standard; peptide; 24 AA.

XX AC

XX ABB40149;

XX DT

XX 04-FEB-2002 (first entry)

XX XX

DE Peptide #7655 encoded by human foetal liver single exon probe.

XX XX

KW Human; foetal liver; gene expression; single exon nucleic acid probe.

XX OS

XX Homo sapiens.

XX PN

WO200157277-A2.

XX PD

XX 09-AUG-2001.

XX XX

XX 30-JAN-2001; 2001WO-US000669.

XX PF

XX 04-FEB-2000; 2000US-0180312P.

XX PR

XX 26-MAY-2000; 2000US-0207456P.

XX PR

XX 30-JUN-2000; 2000US-00608408.

XX PR

XX 03-AUG-2000; 2000US-00632366.

XX PR

XX 21-SEP-2000; 2000US-0234687P.

XX PR

XX 27-SEP-2000; 2000US-0236359P.

XX PR

XX 04-OCT-2000; 2000GB-00024263.

XX PA

(MOLE-) MOLECULAR DYNAMICS INC.

XX XX

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX PI

XX WPI; 2001-483447/52.

XX DR

XX XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human fetal liver.

XX XX

PS Claim 27; SEQ ID NO 32784; 639pp + Sequence Listing; English.

XX XX

CC The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human fetal liver. The present sequence is a peptide encoded by a single exon nucleic acid probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX XX

SQ Sequence 24 AA;

Query Match 33.9%; Score 39; DB 4; Length 24;
Best Local Similarity 50.0%; Pred. No. 67;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 4 FSKTAAGGAAARPG 19
|||

DB 4 FSHVPEANSTAGEKEPG 19
|||

RESULT 4

AAM33792
ID AAM33792 standard; protein; 24 AA.

XX AC

XX AAM33792;

XX XX

```

DT 17-OCT-2001 (first entry)
XX
DE Peptide #7829 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-48897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human placenta.
XX
PS Claim 27; SEQ ID NO 34061; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP:
XX see AA13115-AA15746). The present sequence is a peptide encoded by one
XX such probe. The probes are useful for producing a microarray for
XX predicting, measuring and displaying gene expression in samples derived
XX from human placenta. The probes are useful for antenatal diagnosis of
XX human genetic disorders
XX
XX Sequence 24 AA;
XX
Query Match 33.9%; Score 39; DB 4; Length 24;
Best Local Similarity 50.0%; Pred. No. 67;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 4 FSKTAAGGAAERPG 19
DB ||| ||| |||
4 FSHVPANSTAAGEKPG 19

RESULT 5
AAM73597
ID AAM73597 standard; protein; 24 AA.
XX
AC AAM73597;
XX
XX 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 33903.
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukaemia; lymphoma; myeloma.
XX
XX Homo sapiens.
XX
XX WO200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000668.
XX

```

```

PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-48897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 33903; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is a
XX protein encoded by one of the probes of the invention
XX
XX Sequence 24 AA;
XX
Query Match 33.9%; Score 39; DB 4; Length 24;
Best Local Similarity 50.0%; Pred. No. 67;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 4 FSKTAAGGAAERPG 19
DB ||| ||| |||
4 FSHVPANSTAAGEKPG 19

RESULT 6
AAM60912
ID AAM60912 standard; protein; 24 AA.
XX
AC AAM60912;
XX
XX 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 33017.
XX
XX Human; brain expressed exon; gene expression analysis; probe; microarray;
XX Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000667.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX

```

PT brains.
 XX Example 4; SEQ ID NO 33017; 650pp + Sequence Listing; English.
 XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention
 XX Sequence 24 AA;
 SQ
 Query Match 33.9%; Score 39; DB 4; Length 24;
 Best Local Similarity 50.0%; Pred. No. 67;
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
 QY 4 FSKTAAKGAEAAERPG 19
 ||| | | | | | | |
 DB 4 FSHVPANSTAGEKPG 19
 RESULT 7
 ID ABG55330 standard; peptide; 24 AA.
 AC ABG55330;
 XX
 DT 25-FEB-2003 (first entry)
 XX
 DE Human liver peptide, SEQ ID NO 33978.
 XX
 KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157273-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000664.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488898/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human adult liver.
 XX
 PS Claim 27; SEQ ID NO 33978; 658pp; English.
 XX
 CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult liver.
 CC (I) may be used for predicting, measuring and displaying gene expression
 CC in samples derived from human adult liver. The genes identified may be
 CC involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
 CC associated with coronary heart disease. ABG47348-ABG59930 represent human

CC liver single exon encoded peptides of the invention. Note: The sequence
 CC information for this patent does not appear in the printed specification
 CC but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 24 AA;
 SQ
 Query Match 33.9%; Score 39; DB 4; Length 24;
 Best Local Similarity 50.0%; Pred. No. 67;
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
 QY 4 FSKTAAKGAEAAERPG 19
 ||| | | | | | | |
 DB 4 FSHVPANSTAGEKPG 19
 RESULT 8
 ID ABG43467 standard; peptide; 24 AA.
 AC ABG43467;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 33132.
 XX
 KW Human; single exon probe; asthma; lung cancer; COPD; IID;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000665.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples.
 XX
 PS Claim 27; SEQ ID NO 33132; 634pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of probes
 CC ; the novel set of probes which hybridise at high stringency to a nucleic
 CC acid expressed in the human lung; measuring gene expression in a sample
 CC derived from human lung, comprising (a) contacting the array with a
 CC collection of detectably labeled nucleic acids derived from human lung

CC mRNA, and (b) measuring the label detectably bound to each probe of the
 CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karsenger syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a peptide/protein encoded by a single exon probe of
 CC the invention. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pot_sequences
 CC
 CC Sequence 24 AA;

Query Match 33.9%; Score 39; DB 5; Length 24;
 Best Local Similarity 50.0%; Pred. No. 67;
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 4 FSKTAAGGAAARPG 19
 || | || || ||
 DB 4 FSHVPANSTAGEKPG 19

RESULT 9
 AAW30222
 ID AAW30222 standard; peptide; 16 AA.

AAW30222;

25-MAR-2003 (revised)

21-JAN-1998 (first entry)

KEA16, amphiphilic potential membrane forming peptide.

Membrane-forming peptide; amphiphilic; monovalent metal cation; suture;
 internal lining; slow-diffusion drug delivery vehicle; erythropoietin;
 tissue-type plasminogen activator; haemoglobin; insulin; artificial skin;
 separation matrix; dialysis membranes; filter; collagen; cell migration;
 Alzheimer's disease; scrapie infection; therapy; proteolytic digestion;
 acidic pH resistant; stomach acid; cell adhesion; cell monolayer;
 tissue culture; differentiated cell; stratified cell layer.

Synthetic.

US5670483-A.

23-SEP-1997.

30-NOV-1994; 94US-00345849.

28-DEC-1992; 92US-00973326.

(MASI) MASSACHUSETTS INST TECHNOLOGY.

Rich A, Holmes T, Zhang S, Lockshin C;

XX

DR WPI; 1997-479506/44.
 XX
 PT Membranes formed by self-assembly of amphiphilic peptide(s) - useful as
 bio-material(s), separation matrices, drug delivery vehicles, etc.
 XX
 PS Claim 30; Col 63; 49pp; English.
 XX

AAW30219-W30252 are amphiphilic peptides used in the macroscopic membrane
 (MM) of the invention. The MM is formed by self-assembly of the
 amphiphilic peptides (with alternating hydrophilic and hydrophobic amino
 acids) in an aqueous medium containing monovalent metal cations. The MM
 are stable in serum, are non-cytotoxic, and are useful in biomaterial
 applications, such as medical products (e.g. sutures), or internal
 linings. The MM are useful as slow-diffusion drug delivery vehicles for
 protein-type drugs, including erythropoietin, tissue-type plasminogen
 activator, synthetic haemoglobin and insulin. They can be used in
 applications, such as separation matrices (e.g., dialysis membranes). The
 extremely small pore size (interfilament distance) of the MM makes them
 useful as filters. Collagen may be combined with the peptides to produce
 membranes for use as artificial skin. The MM may be used for making very
 thin, transparent fabric. Drugs which inhibit the self assembly of the
 peptides into filaments or filamentous membranes may be useful for
 treating Alzheimer's disease or scrapie infection. As they are resistant
 to proteolytic digestion and alkaline and acidic pH (such as stomach
 acid), drug delivery vehicles made of the MM could be taken orally. The
 charged residues and conformation of the MM promote cell adhesion and
 migration. The permeability of the MM also permits diffusion of small
 molecules, to the underside of cell monolayers, useful for tissue culture
 of differentiated cells and/or stratified cell layers. (Updated on 25-MAR
 -2003 to correct PF field.)
 XX
 SQ Sequence 16 AA;

Query Match 32.2%; Score 37; DB 2; Length 16;
 Best Local Similarity 56.2%; Pred. No. 89;
 Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 6 KTAAGGAAARPGEA 21

1 KAEAKAEAKAEAKAE 16

RESULT 10

AAW32964

ID AAY32964 standard; peptide; 16 AA.

AC AAY32964;

09-NOV-1999 (first entry)

Membrane forming amphiphilic peptide KEA16.

Membrane forming peptide; cell culturing; macroscopic membrane;
 amphiphilic peptide; slow-diffusion drug delivery system; cell growth;
 artificial skin; separation matrix; artificial tissue; scrapie infection;
 Alzheimer's disease; liver cirrhosis; kidney amyloidosis;
 protein conformational disease.

Synthetic.

US5955343-A.

21-SEP-1999.

22-AUG-1994; 94US-00293284.

28-DEC-1992; 92US-00973326.

(MASI) MASSACHUSETTS INST TECHNOLOGY.

Rich A, Dipersio CM, Lockshin C, Holmes T, Zhang S;

WPI; 1999-539576/45.

XX

XX Cell cultures utilizing stable macroscopic membranes formed by the self-
 PT assembly of amphiphilic peptides.
 XX
 XX Example 5; Col 26; 49pp; English.
 PS
 XX This sequence represents a membrane forming amphiphilic peptide used in
 CC the in vitro method of the invention. The method is for culturing cells
 CC utilising stable macroscopic membranes formed by the self-assembly of
 CC amphiphilic peptides. The peptides are salt-induced to form insoluble and
 CC protease-resistant protein filaments with a beta-sheet secondary
 CC structure. The membranes may be useful in a wide variety of
 CC medical, research, industrial and biomaterial applications such as slow-
 CC diffusion drug delivery systems, artificial skin and separation matrices.
 CC The membranes may be used to support in vitro cell attachment and growth
 CC and for supporting artificial tissue (e.g. for in vivo use as implants).
 CC They are particularly useful as experimental models for Alzheimer's
 CC disease and scrapie infection and so may be used in disease modelling
 CC experiments and to assay for agents which modulate the disease processes.
 CC Additionally, they may be used in this way to study liver cirrhosis,
 CC kidney amyloidosis and other protein conformational diseases. The
 CC membrane in the membrane/cell mixture produced in the method: (i)
 CC supports cellular attachment and growth; (ii) self-assembles to form
 CC large, macroscopic membranes that are insoluble and stable in aqueous
 CC solutions, serum and ethanol; (iii) is highly resistant to heat,
 CC alkaline/acidic pHs, chemical denaturants and proteolytic digestion; (iv)
 CC is non-cytotoxic and non-immunogenic; (v) is visible to the naked eye if
 CC dyed but is otherwise transparent; (vi) may form thin, permeable, high
 CC density sheets or fibril like structures with simple structures, high
 CC tensile strength and a porous structure; (vii) may be metabolised by
 CC humans and animals; (viii) is inexpensive to produce; and (ix) can be
 CC produced and stored in sterile conditions
 XX
 SQ Sequence 16 AA;

Query Match 32.2%; Score 37; DB 2; Length 16;
 Best Local Similarity 56.2%; Pred. No. 89;
 Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 6 KTAAGGEAAARPGEA 21
 |||||
 Db 1 KAEAKAEAKAEAKAEA 16

RESULT 11
 AAE38860
 ID AAE38860 standard; peptide; 16 AA.
 XX
 AC AAE38860;

DE 18-DEC-2003 (first entry)
 XX Membrane forming peptide, KEA16.
 DE
 XX Macroscopic membrane; biomaterial; medical product; internal lining;
 KW artificial skin; slow-diffusion drug delivery; chromatography column;
 KW separation matrix; suture; dialysis membrane; viral filter; fabric;
 KW pathology; Alzheimer's disease; scrapie infection.
 XX
 OS Unidentified.
 XX
 PN US6548630-B1.
 XX
 PD 15-APR-2003.
 XX
 XX 22-JUL-1997; 97US-00898300.
 PF
 XX 28-DEC-1992; 92US-00973326.
 PR
 XX 30-NOV-1994; 94US-00346849.
 XX
 PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
 XX
 PI Zhang S, Lockshin C, Rich A, Holmes T;

XX WPI; 2003-584339/55.
 DR
 XX Macroscopic membrane useful in biomaterial applications e.g. sutures is
 PT formed by self-assembly of amphiphilic peptides in an aqueous solution
 PT containing monovalent metal cations.
 XX
 XX Example 5; Col 25; 51pp; English.
 PS
 XX The invention relates to a macroscopic membrane formed by self-assembly
 CC of amphiphilic peptides in an aqueous solution containing monovalent
 CC metal cations. The peptides have alternating hydrophobic and hydrophilic
 CC amino acids and are complementary and structurally compatible. The
 CC membranes are useful as biomaterial for medical products (e.g. sutures,
 CC artificial skin, internal linings), as vehicles for slow-diffusion drug
 CC delivery (preferably for protein type drugs e.g. erythropoietin, tissue
 CC type plasminogen activator, synthetic haemoglobin and insulin), as
 CC separation matrices (e.g. dialysis membranes, chromatography columns),
 CC filters for the removal of viruses and other microscopic contaminants),
 CC for other uses requiring permeable and water-insoluble material (e.g. for
 CC culturing cell monolayers including differentiated cells and/or
 CC stratified cell layers), for the preparation of very thin and transparent
 CC fabric and as a model system for investigating the properties of
 CC biological protein structures and providing insights into the pathology
 CC and potential treatment of conditions involving the presence of the
 CC proteins and proteinaceous structures (e.g. Alzheimer's disease, scrapie
 CC infection and in origin of life studies related to cell membranes and
 CC cellular compartmentalisation). The present sequence is a membrane
 CC forming peptide used in the invention
 XX

SQ Sequence 16 AA;

Query Match 32.2%; Score 37; DB 7; Length 16;

Best Local Similarity 56.2%; Pred. No. 89;
 Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 6 KTAAGGEAAARPGEA 21
 |||||
 Db 1 KAEAKAEAKAEAKAEA 16

RESULT 12
 AAY20179
 ID AAY20179 standard; protein; 19 AA.
 XX
 AC AAY20179;

XX 22-JUL-1999 (first entry)
 DT
 XX Human beta-amyloid precursor protein mutant fragment 4.

DE Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX WO9845322-A2.
 PN
 XX 15-OCT-1998.
 PD
 XX 02-APR-1998; 98WO-IB000705.
 PF
 XX 10-APR-1997; 97US-0043163P.
 PR
 XX

PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 PA (UYUT-) RIJKSUNIV UTRECHT.

PI Van Leeuwen FW, Grosveld FG, Burbach JPH;

XX WPI; 1998-609901/51.

DR N-PSDB; AAX75753.

XX Diagnosing disease by detecting frameshift mutations in RNA or

PT corresponding protein mutations - used to diagnose cancer and

PT neurological diseases, particularly Alzheimer's disease, and also for

PT treatment and prevention with specific ribozymes or wild-type RNA.

XX Disclosure; Fig 2; 258pp; English.

CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin 1, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A

XX Sequence 19 AA;

Query Match 32.2%; Score 37; DB 2; Length 19;

Best Local Similarity 50.0%; Pred. No. 1.1e+02;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 SKTAAGGEAAERP 18

Db 2 SRTGASGAASARP 15

RESULT 13

ADB81153
 ID ADB81153 standard; peptide; 24 AA.

XX ADB81153;

DT 04-DEC-2003 (first entry)

DE Human pancreatic/heart tissue quantified by LC-MS analysis, SEQ ID No 44.

XX quantitating relative abundance; liquid phase tandem mass spectroscopy;
 KW data set; peptide profile; toxicology analysis; classification; human.

XX Homo.

OS sapiens.

XX W0200297703-A2.

XX 05-DEC-2002.

XX 30-MAY-2002; 2002WO-CA000801.

XX 30-MAY-2001; 2001CA-02349265.

XX (EMILI/) EMILI A.

PA (CAGN/) CAGNEY G.

XX

PI Emili A, Cagney G;

XX WPI; 2003-175129/17.

XX Identifying and quantifying proteins/peptide sequences, comprises

PT identifying peptides by liquid phase tandem mass spectroscopy sequencing

PT and compiling data set/peptide profile containing collection of peptide

PT sequences.

XX Example 3; Page 49; 84pp; English.

XX The invention relates to a novel method for identifying constituent
 CC proteins/peptide sequences or quantitating relative abundance of proteins
 CC in two samples, for a cell type, tissue or pathological sample using a
 CC database having peptide profile libraries with multiple protein/peptide
 CC sequences. The method comprises identifying peptide species by liquid
 CC phase tandem mass spectroscopy sequencing and compiling a data
 CC set/peptide profile. The novel method is useful in toxicology analysis.

CC In the method of the invention, there is no requirement for prior
 CC knowledge about the functions of the responsive peptides or parental
 CC proteins. Protein functions deduced from comparisons of profiles in a
 CC database can be derived from very subtle physiological responses. For
 CC instance, even though peptide levels may change only slightly in response
 CC to an experimental treatment, coordinate changes among many measured
 CC peptide abundances can be sufficient to characterise that phenotype. The
 CC large number of peptides measured make it unlikely that an unrelated
 CC physiological state will have an identical profile, even though this may
 CC not be apparent when using conventional experiments that measure the
 CC levels of one or a few proteins. Closely related profiles can be classed
 CC together, thus improving the understanding of the underlying biological
 CC basis of the classifications. This sequence represents a peptide
 CC sequenced and quantified from a human pancreatic or heart tissue sample
 CC in a single liquid chromatography-mass spectrometry (LC-MS) analysis.
 CC used in the method for the measurement of protein relative abundance in
 CC complex mixtures.

XX Sequence 24 AA;

Query Match 32.2%; Score 37; DB 7; Length 24;

Best Local Similarity 47.6%; Pred. No. 1.4e+02;

Matches 10; Conservative 2; Mismatches 7; Indels 2; Gaps 1;

QY 6 KTAAGGEAAERP--GEAAVA 24

Db 1 KAASLGSSQPSRPHVGEAATA 21

RESULT 14

AAY36444

ID AAY36444 standard; protein; 20 AA.

XX AAY36444;

DT 17-SEP-1999 (first entry)

DE Fragment of human secreted protein encoded by gene 11.

XX Human; secreted protein; cancer; tumour; developmental abnormality;
 KW foetal deficiency; blood disorder; immune system disorder; inflammation;
 KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
 KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
 KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
 KW digestive disorder; endocrine disorder; infection; AIDS.

XX Homo sapiens.

XX W09931117-A1.

XX 24-JUN-1999.

XX 17-DEC-1998; 98WO-US027059.

XX 18-DEC-1997; 97US-0068006P.

```

PR 18-DEC-1997; 97US-0068007P.
PR 18-DEC-1997; 97US-0068008P.
PR 18-DEC-1997; 97US-00680053P.
PR 18-DEC-1997; 97US-00680054P.
PR 18-DEC-1997; 97US-00680057P.
PR 18-DEC-1997; 97US-00680064P.
PR 18-DEC-1997; 97US-0070923P.
PR 19-DEC-1997; 97US-0068169P.
PR 19-DEC-1997; 97US-0068365P.
PR 19-DEC-1997; 97US-0068367P.
PR 19-DEC-1997; 97US-0068368P.
PR 19-DEC-1997; 97US-0068369P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Moore PA, Ruben SM, Carter KC, Shi Y, Rosen CA, Soppet DR;
PI Kyaw H, Wei Y, Florence K, Duan RD, Florence C, Greene JM, Feng P;
PI Ferrie AM, Yu G, Janat F, Ni J;
XX
XX WPI; 1999-418749/35.
XX
XX New isolated human genes encoding secreted polypeptides.
XX
XX Disclosure; Page 450; 537pp; English.
XX
XX AAX97916 to AAX98029 represent 110 isolated human secreted protein genes.
XX AAX36224 to AAX36727 represent the secreted proteins encoded by the 110
XX human genes. The genes and their corresponding secreted polypeptides are
XX useful for preventing, treating or ameliorating medical conditions, e.g.
XX by protein or gene therapy. Also pathological conditions can be diagnosed
XX by determining the amount of the new polypeptides in a sample or by
XX determining the presence of mutations in the new genes. Specific uses are
XX described for each of the 110 genes, based on which tissues they are most
XX highly expressed in, and include developing products for the diagnosis or
XX treatment of cancer, tumours, developmental abnormalities and foetal
XX deficiencies, blood disorders, diseases of the immune system, autoimmune
XX diseases, inflammation, allergies, Alzheimer's and cognitive disorders,
XX schizophrenia, arthritis, asthma, psoriasis, sepsis, skin disorders,
XX atherosclerosis, diabetes, cardiovascular disorders, kidney disorders,
XX digestive/endocrine disorders, infections and AIDS. The polypeptides are
XX also useful for identifying their binding partners. The sequences given
XX in AAX97907 to AAX97915 and AAX36223 are used in the exemplification of
XX the present invention
XX
XX Sequence 20 AA;
XX
XX Query Match 31.3%; Score 36; DB 2; Length 20;
XX Best Local Similarity 70.0%; Pred. No. 1.6e+02;
XX Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 13 AAAERPGEEA 22
XX Db 8 AQAEKPGEGA 17
XX
XX RESULT 15
XX ADA11979
XX ID ADA11979 standard; protein; 20 AA.
XX
XX AC ADA11979;
XX
XX DT 06-NOV-2003 (first entry)
XX
XX DE Human novel secreted protein associated polypeptide #249.
XX
XX KW cancer; inflammation; immune disorder; neurological disorder;
XX blood clotting disorder; food additive; food preservative;
XX storage capability; fat content; nutritional component; human;
XX secreted protein.
XX
XX OS Homo sapiens.
XX
XX PN US2003055236-A1.

```

```

XX 20-MAR-2003.
XX
XX PF 14-MAR-2002; 2002US-00097065.
XX
XX 18-DEC-1997; 97US-0068006P.
XX 18-DEC-1997; 97US-0068007P.
XX 18-DEC-1997; 97US-0068008P.
XX 18-DEC-1997; 97US-00680053P.
XX 18-DEC-1997; 97US-00680054P.
XX 18-DEC-1997; 97US-00680057P.
XX 18-DEC-1997; 97US-00680064P.
XX 18-DEC-1997; 97US-0070923P.
XX 19-DEC-1997; 97US-0068169P.
XX 19-DEC-1997; 97US-0068365P.
XX 19-DEC-1997; 97US-0068367P.
XX 19-DEC-1997; 97US-0068368P.
XX 19-DEC-1997; 97US-0068369P.
XX 17-DEC-1998; 98WO-US027059.
XX 17-JUN-1999; 99US-00334595.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Moore PA, Ruben SM, Carter KC, Shi Y, Rosen CA, Soppet DR;
PI Kyaw H, Wei Y, Florence KA, Duan DR, Florence C, Greene JM, Feng P;
PI Ferrie AM, Yu G, Janat F, Ni J;
XX
XX WPI; 2003-567105/53.
XX
XX New secreted HKABT24 nucleic acid molecules and polypeptides, useful for
XX preventing, treating, or ameliorating a medical condition, such as
XX cancer, inflammation, immune disorders, neurological and blood clotting
XX disorders.
XX
XX Disclosure; Page 34; 118pp; English.
XX
XX The invention relates to an isolated HKABT24 nucleic acid molecule. The
XX polypeptides, nucleic acids and antibodies are useful for diagnosing a
XX pathological condition or a susceptibility to a pathological condition,
XX for preventing, treating, or ameliorating a medical condition, such as
XX cancer, inflammation and other immune disorders, neurological and blood
XX clotting disorders. The nucleic acids are also useful for chromosome
XX identification, radiation hybrid mapping or long-range restriction
XX mapping. The polypeptides and antibodies are useful for providing
XX immunological probes for differential identification of the tissues
XX immunohistochemistry assays. The polypeptide, polynucleotide, agonist or
XX antagonist may also be used as a food additive or preservative to
XX increase or decrease storage capabilities, fat content or other
XX nutritional components. The present sequence represents the amino acid
XX sequence of a novel human secreted protein associated polypeptide. Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from USPTO
XX at seqdata.uspto.gov.uk/sequence.html?DocID=20030055236.
XX
XX Sequence 20 AA;
XX
XX Query Match 31.3%; Score 36; DB 6; Length 20;
XX Best Local Similarity 70.0%; Pred. No. 1.6e+02;
XX Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 13 AAAERPGEEA 22
XX Db 8 AQAEKPGEGA 17
XX
XX Search completed: October 7, 2004, 17:53:30
XX Job time : 88.5 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:54:08 ; Search time 83.5 Seconds
(without alignments)
92.493 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GAQFSKTAAKGEAAERPGAAVA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 284360

Minimum DB seq length: 0

Maximum DB seq length: 24

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	100.0	24	12	US-10-180-753-1
2	115	100.0	24	16	US-10-802-644-1
3	39	33.9	24	9	US-09-864-761-44320
4	37	32.2	16	16	US-10-390-472-35
5	37	32.2	20	10	US-09-939-537-52
6	36	31.3	20	12	US-10-372-876-507
7	36	31.3	20	14	US-10-097-065-507
8	35	30.4	21	16	US-10-109-048-902
9	35	30.4	21	16	US-10-109-048-904
10	34	29.6	16	16	US-10-390-472-33
11	33	28.7	15	9	US-09-765-301-28
12	33	28.7	15	10	US-09-765-644A-28
13	33	28.7	15	15	US-10-438-538-32
14	33	28.7	16	9	US-09-778-200-6
15	33	28.7	16	9	US-09-778-200-21

16	33	28.7	16	14	US-10-192-832-6	Sequence 6, Appli
17	33	28.7	16	14	US-10-192-832-22	Sequence 22, Appl
18	33	28.7	16	16	US-10-390-472-34	Sequence 34, Appl
19	33	28.7	16	16	US-10-390-472-36	Sequence 36, Appl
20	32	27.8	16	9	US-09-778-200-2	Sequence 2, Appli
21	32	27.8	16	14	US-10-192-832-2	Sequence 2, Appli
22	32	27.8	20	12	US-10-283-017-2058	Sequence 2058, Ap
23	32	27.8	21	10	US-09-974-879-413	Sequence 413, App
24	32	27.8	21	10	US-09-305-736-413	Sequence 413, App
25	32	27.8	21	11	US-09-818-683-413	Sequence 413, App
26	32	27.8	21	12	US-10-621-401-413	Sequence 413, App
27	32	27.8	21	12	US-09-818-683-413	Sequence 413, App
28	32	27.8	21	15	US-10-448-163-28	Sequence 28, Appl
29	32	27.8	22	12	US-10-619-778-7	Sequence 7, Appli
30	32	27.8	22	12	US-09-726-792A-7	Sequence 7, Appli
31	32	27.8	22	15	US-10-462-138-6	Sequence 6, Appli
32	32	27.8	22	16	US-10-755-784-7	Sequence 7, Appli
33	31.5	27.4	20	10	US-09-895-298-150	Sequence 150, App
34	31	27.0	15	9	US-09-765-301-10	Sequence 10, Appl
35	31	27.0	15	9	US-09-765-301-24	Sequence 24, Appl
36	31	27.0	15	10	US-09-765-644A-10	Sequence 10, Appl
37	31	27.0	15	10	US-09-765-644A-24	Sequence 24, Appl
38	31	27.0	15	15	US-10-438-538-14	Sequence 14, Appl
39	31	27.0	15	15	US-10-438-538-28	Sequence 28, Appl
40	31	27.0	16	10	US-09-847-586-6	Sequence 6, Appli
41	31	27.0	16	10	US-09-847-586-14	Sequence 14, Appl
42	31	27.0	16	10	US-09-847-586-15	Sequence 15, Appl
43	31	27.0	16	10	US-09-847-586-16	Sequence 16, Appl
44	31	27.0	16	10	US-09-847-586-17	Sequence 17, Appl
45	31	27.0	16	10	US-09-847-586-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-10-180-753-1
; Sequence 1, Application US/10180753
; Publication No. US20030013652A1
; GENERAL INFORMATION:
; APPLICANT: Martin, Linda
; APPLICANT: Adler, Kenneth
; APPLICANT: Li, Yuehua
; TITLE OF INVENTION: BLOCKING PEPTIDE FOR INFLAMMATORY CELL SECRETION
; FILE REFERENCE: 5051.574
; CURRENT APPLICATION NUMBER: US/10/180,753
; CURRENT FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: N-terminal myristoylation
US-10-180-753-1

Query Match 100.0%; Score 115; DB 12; Length 24;
Best Local Similarity 100.0%; Pred. No. 8.6e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAAKGEAAERPGAAVA 24
|||||
Db 1 GAQFSKTAAKGEAAERPGAAVA 24

RESULT 2
US-10-802-644-1
; Sequence 1, Application US/10802644

;; PRIOR FILING DATE: 1998-12-17
;; PRIOR APPLICATION NUMBER: 60/070,923
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,007
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,057
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,006
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,369
;; PRIOR FILING DATE: 1997-12-19
;; PRIOR APPLICATION NUMBER: 60/068,367
;; PRIOR FILING DATE: 1997-12-19
;; PRIOR APPLICATION NUMBER: 60/068,368
;; PRIOR FILING DATE: 1997-12-19
;; PRIOR APPLICATION NUMBER: 60/068,169
;; PRIOR FILING DATE: 1997-12-19
;; PRIOR APPLICATION NUMBER: 60/068,053
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,064
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,054
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,008
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,365
;; PRIOR FILING DATE: 1997-12-19
;; NUMBER OF SEQ ID NOS: 672
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 507
;; LENGTH: 20
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-097-065-507

Query Match 31.3%; Score 36; DB 14; Length 20;
Best Local Similarity 70.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 13 AAARPGEAA 22
|||:|||||
DB 8 AQAEPGEA 17

RESULT 8
US-10-109-048-902
; Sequence 902, Application US/10109048
; Publication No. US20040107461A1
; GENERAL INFORMATION:
; APPLICANT: COMMURI, PADMA
; APPLICANT: KEELING, PETER L.
; APPLICANT: RAMIREZ, NONA
; APPLICANT: MCKEAN, ANGELA
; APPLICANT: GAO, ZHONG
; APPLICANT: GUAN, HANPING
; TITLE OF INVENTION: GLUCAN CHAIN LENGTH DOMAINS
; FILE REFERENCE: 2461-76
; CURRENT APPLICATION NUMBER: US/10/109,048
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: 60/279,720
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 1154
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 902
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Accession No. 15903076
US-10-109-048-902

Query Match 30.4%; Score 35; DB 16; Length 21;
Best Local Similarity 58.3%; Pred. No. 3.3e+02;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 GAQFSKTAAKGE 12
|||:|||||
DB 10 GAQFSKTGGLGD 21
RESULT 9
US-10-109-048-904
; Sequence 904, Application US/10109048
; Publication No. US20040107461A1
; GENERAL INFORMATION:
; APPLICANT: COMMURI, PADMA
; APPLICANT: KEELING, PETER L.
; APPLICANT: RAMIREZ, NONA
; APPLICANT: MCKEAN, ANGELA
; APPLICANT: GAO, ZHONG
; APPLICANT: GUAN, HANPING
; TITLE OF INVENTION: GLUCAN CHAIN LENGTH DOMAINS
; FILE REFERENCE: 2461-76
; CURRENT APPLICATION NUMBER: US/10/109,048
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: 60/279,720
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 1154
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 904
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Accession No. 15900991
US-10-109-048-904

Query Match 30.4%; Score 35; DB 16; Length 21;
Best Local Similarity 58.3%; Pred. No. 3.3e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 GAQFSKTAAKGE 12
|||:|||||
DB 10 GAQFSKTGGLGD 21

RESULT 10
US-10-390-472-33
; Sequence 33, Application US/10390472
; Publication No. US20040087013A1
; GENERAL INFORMATION:
; APPLICANT: Holmes, Todd
; Zhang, Shuguang
; Rich, Alexander
; DiPersio, C. Michael
; Lockshin, Curtis
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
; SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
; THEREFOR
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESS: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/390,472
; FILING DATE: 17-Mar-2003
; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/293,284
;; FILING DATE: 22-AUG-1994
;; APPLICATION NUMBER: 07/973,326
;; FILING DATE: 28-DEC-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Brook, David E.
;; REGISTRATION NUMBER: 22,592
;; REFERENCE/DOCKET NUMBER: MIT-6008A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617) 861-6240
;; TELEFAX: (617) 861-9540
;; INFORMATION FOR SEQ ID NO: 33:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 16 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-10-390-472-33

Query Match 29.6%; Score 34; DB 16; Length 16;
Best Local Similarity 50.0%; Pred. No. 3.5e+02;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 SKTAAGKEAAERPG 20
Db 1 AKAAKAAKAAKAAE 16
:::| | | | |

RESULT 11
US-09-765-301-28
; Sequence 28, Application US/09765301
; Patent No. US20020037848A1
; GENERAL INFORMATION:
; APPLICANT: EISENBACH-SCHWARTZ, Michal
; APPLICANT: YOLLES, Etti
; TITLE OF INVENTION: THE USE OF COPOLYMER 1 AND RELATED PEPTIDES AND POLYPEPTIDES AND
; FILE REFERENCE: EIS-SCHWARTZ18
; CURRENT APPLICATION NUMBER: US/09/765,301
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 06/209,799
; PRIOR FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: 09/620,216
; PRIOR FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-09-765-301-28

Query Match 28.7%; Score 33; DB 9; Length 15;
Best Local Similarity 50.0%; Pred. No. 4.5e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 AOFSKTAAGKEAAA 15
Db 1 AEYAKAAAAA 14
:::| | | | |

RESULT 12
US-09-765-644A-28
; Sequence 28, Application US/09765644A
; Publication No. US20030004099A1
; GENERAL INFORMATION:
; APPLICANT: EISENBACH-SCHWARTZ, Michal
; APPLICANT: COHEN, Irun R.

;; APPLICANT: SELA, Michael
;; APPLICANT: YOLLES, Etti
;; APPLICANT: KIPNIS, Jonathan
;; TITLE OF INVENTION: THE USE OF COPOLYMER 1 AND RELATED PEPTIDES AND POLYPEPTIDES AND
;; FILE REFERENCE: EIS-SCHWARTZ13B
;; CURRENT APPLICATION NUMBER: US/09/765,644A
;; CURRENT FILING DATE: 2001-01-22
;; PRIOR APPLICATION NUMBER: 09/487,793
;; PRIOR FILING DATE: 2000-01-20
;; PRIOR APPLICATION NUMBER: 06/209,799
;; PRIOR FILING DATE: 2000-06-07
;; PRIOR APPLICATION NUMBER: 09/620,216
;; PRIOR FILING DATE: 2000-07-20
;; NUMBER OF SEQ ID NOS: 33
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 28
;; LENGTH: 15
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Synthetic construct
US-09-765-644A-28

Query Match 28.7%; Score 33; DB 10; Length 15;
Best Local Similarity 50.0%; Pred. No. 4.5e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 AOFSKTAAGKEAAA 15
Db 1 AEYAKAAAAA 14
:::| | | | |

RESULT 13
US-10-438-538-32
; Sequence 32, Application US/10438538
; Publication No. US20040006022A1
; GENERAL INFORMATION:
; APPLICANT: Strominger, Jack L.
; APPLICANT: Fridkis-Hareli, Masha
; TITLE OF INVENTION: Synthetic Peptides and Methods of use for Autoimmune
; FILE REFERENCE: 24655-013DIV2
; CURRENT APPLICATION NUMBER: US/10/438,538
; CURRENT FILING DATE: 2003-05-15
; PRIOR APPLICATION NUMBER: 09/359,099
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 60/093,859
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: 60/123,675
; PRIOR FILING DATE: 1999-03-09
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide of predetermined sequence for testing of
; OTHER INFORMATION: activity in MHC Class II assays.
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(15)
US-10-438-538-32

Query Match 28.7%; Score 33; DB 15; Length 15;
Best Local Similarity 50.0%; Pred. No. 4.5e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 AOFSKTAAGKEAAA 15
Db 1 AEYAKAAAAA 14
:::| | | | |

RESULT 14
US-09-778-200-6
; Sequence 6, Application US/09778200
; Patent No. US20020160471A1
; GENERAL INFORMATION:
; APPLICANT: Kisdai, John
; APPLICANT: Grodzinsky, Alan
; APPLICANT: Zhang, Shuguang
; TITLE OF INVENTION: Peptide Scaffold Encapsulation of Tissue
; TITLE OF INVENTION: Cells and Uses Thereof
; FILE REFERENCE: 01997/537001
; CURRENT APPLICATION NUMBER: US/09/778,200
; CURRENT FILING DATE: 2000-02-06
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Designed Peptide
US-09-778-200-6

Query Match 28.7%; Score 33; DB 9; Length 16;
Best Local Similarity 61.5%; Pred. No. 4.8e+02;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 AKGEAAERPGEA 21
|| || || || ||
Db 3 AKAEAKAEAEAE 15

RESULT 15
US-09-778-200-21
; Sequence 21, Application US/09778200
; Patent No. US20020160471A1
; GENERAL INFORMATION:
; APPLICANT: Kisdai, John
; APPLICANT: Grodzinsky, Alan
; APPLICANT: Zhang, Shuguang
; TITLE OF INVENTION: Peptide Scaffold Encapsulation of Tissue
; TITLE OF INVENTION: Cells and Uses Thereof
; FILE REFERENCE: 01997/537001
; CURRENT APPLICATION NUMBER: US/09/778,200
; CURRENT FILING DATE: 2000-02-06
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Designed Peptide
US-09-778-200-21

Query Match 28.7%; Score 33; DB 9; Length 16;
Best Local Similarity 50.0%; Pred. No. 4.8e+02;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 KTAKEGAAERPGEA 21
| | | | | | | |
Db 1 KAKAKAEAEAEAE 16

Search completed: October 7, 2004, 18:10:46
Job time : 85.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:50:43 ; Search time 22 Seconds
(without alignments)
56.319 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GAQFSKTAAGBAARPGAAVA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 189146

Minimum DB seq length: 0

Maximum DB seq length: 24

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	32.2	16	1	US-08-346-849-35
2	37	32.2	16	2	US-08-293-284A-35
3	37	32.2	16	4	US-08-898-300-35
4	37	32.2	19	2	US-08-726-306A-72
5	37	32.2	20	2	US-08-284-391B-52
6	37	32.2	20	3	US-09-218-950-52
7	34	29.6	16	1	US-08-346-849-33
8	34	29.6	16	2	US-08-293-284A-33
9	34	29.6	16	4	US-08-898-300-33
10	34	29.6	21	2	US-08-557-309B-58
11	34	29.6	21	2	US-08-929-414-11
12	34	29.6	21	3	US-08-834-306-56
13	34	29.6	21	3	US-08-834-306-58
14	34	29.6	21	3	US-08-993-674A-56
15	34	29.6	21	3	US-08-993-674A-58
16	34	29.6	21	4	US-09-256-976-56
17	34	29.6	21	4	US-09-256-976-58
18	33	28.7	16	1	US-08-346-849-34
19	33	28.7	16	1	US-08-346-849-36
20	33	28.7	16	2	US-08-293-284A-34
21	33	28.7	16	2	US-08-293-284A-36
22	33	28.7	16	4	US-08-898-300-34
23	33	28.7	16	4	US-08-898-300-36
24	32	27.8	21	1	US-08-127-499A-33
25	32	27.8	21	1	US-08-482-847-33
26	32	27.8	22	3	US-08-990-015-5
27	32	27.8	22	4	US-09-589-978C-7

28	31	27.0	17	4	US-09-140-749-14	Sequence 14, Appl
29	31	27.0	18	4	US-09-005-167A-99	Sequence 99, Appl
30	31	27.0	21	2	US-08-997-080-6	Sequence 6, Appl
31	31	27.0	21	2	US-08-997-362-6	Sequence 6, Appl
32	31	27.0	21	3	US-08-873-970-6	Sequence 6, Appl
33	31	27.0	21	3	US-09-095-855-6	Sequence 6, Appl
34	31	27.0	21	3	US-08-705-347A-6	Sequence 6, Appl
35	31	27.0	21	3	US-09-143-124-22	Sequence 22, Appl
36	31	27.0	21	4	US-09-324-542-6	Sequence 6, Appl
37	31	27.0	21	4	US-09-205-426-6	Sequence 6, Appl
38	31	27.0	21	4	US-09-200-643-6	Sequence 6, Appl
39	31	27.0	22	4	US-07-946-180B-16	Sequence 16, Appl
40	30	26.1	16	1	US-08-346-849-43	Sequence 43, Appl
41	30	26.1	16	2	US-08-293-284A-43	Sequence 43, Appl
42	30	26.1	16	4	US-09-230-548-17	Sequence 17, Appl
43	30	26.1	16	4	US-08-898-300-43	Sequence 43, Appl
44	30	26.1	20	2	US-08-747-137-61	Sequence 61, Appl
45	30	26.1	21	2	US-08-557-309B-60	Sequence 60, Appl

ALIGNMENTS

RESULT 1
US-08-346-849-35
; Sequence 35, Application US/08346849
; Patent No. 5670483
; GENERAL INFORMATION:
; APPLICANT: Zhang, Shuguang
; APPLICANT: Lockshin, Curtis
; APPLICANT: Rich, Alexander
; APPLICANT: Holmes, Todd
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
; TITLE OF INVENTION: THEREFOR
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/346,849
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,326
; FILING DATE: 28 DECEMBER 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: MIT-6008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-346-849-35

Query Match 32.2%; Score 37; DB 1; Length 16;
Best Local Similarity 56.2%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
QY      6 KTAAKGEAAARPGEA 21
      |||||
Db      1 KAEAKAEAKAEAKAEA 16

RESULT 2
US-08-293-284A-35
; Sequence 35, Application US/08293284A
; Patent No. 5955343
; GENERAL INFORMATION:
; APPLICANT: Holmes, Todd
; APPLICANT: Zhang, Shuguang
; APPLICANT: Rich, Alexander
; APPLICANT: DiPersio, C. Michael
; APPLICANT: Lockshin, Curtis
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/293,284A
; FILING DATE: 22-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: 07/973,326
; FILING DATE: 28-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: MIT-6008A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-293-284A-35

Query Match      32.2%; Score 37; DB 2; Length 16;
Best Local Similarity 56.2%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      6 KTAAKGEAAARPGEA 21
      |||||
Db      1 KAEAKAEAKAEAKAEA 16

RESULT 3
US-08-898-300-35
; Sequence 35, Application US/08898300
; Patent No. 6548630
; GENERAL INFORMATION:
; APPLICANT: Zhang, Shuguang
; APPLICANT: Lockshin, Curtis
; APPLICANT: Rich, Alexander
; APPLICANT: Holmes, Todd
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY

; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/898,300
; FILING DATE: 22 JULY 1997
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER: 08/346,849
; FILING DATE: 30 NOVEMBER 1994
; APPLICATION NUMBER: 07/973,326
; FILING DATE: 28 DECEMBER 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: MIT-6008FB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 861-6240
; TELEFAX: (781) 861-9540
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-898-300-35

Query Match      32.2%; Score 37; DB 4; Length 16;
Best Local Similarity 56.2%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      6 KTAAKGEAAARPGEA 21
      |||||
Db      1 KAEAKAEAKAEAKAEA 16

RESULT 4
US-08-726-306A-72
; Sequence 72, Application US/08726306A
; Patent No. 5958684
; GENERAL INFORMATION:
; APPLICANT: van Leeuwen, Frederik Willem
; APPLICANT: Burbach, Johannes Peter Henri
; APPLICANT: Grosveld, Franklin G.
; TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
; NUMBER OF SEQUENCES: 189
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1 Financial Center
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,306A
```


; FILING DATE: 02-Oct-1996
; PRIOR APPLICATION DATA: GB 95/20080.4
; APPLICATION NUMBER: 02-Oct-1995
; FILING DATE: 02-Oct-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,832
; FILING DATE: 01-Jan-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Ph.D., Kathleen M.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
; TELEPHONE: (617) 345-9100
; TELEFAX: (617) 345-9111
; INFORMATION FOR SEQ ID NO: 72:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-726-306A-72

Query Match 32.2%; Score 37; DB 2; Length 19;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 SKTAAKGEAAERPE 18
|:|:|:|:|:|:|
Db 2 SRTGASGAASSARP 15

RESULT 5
US-08-284-391B-52
; Sequence 52, Application US/08284391B
; Patent No. 5851828
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,391B
; FILING DATE: 02-AUG-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200

; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-284-391B-52

Query Match 32.2%; Score 37; DB 2; Length 20;
Best Local Similarity 56.2%; Pred. No. 22;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 6 KTAAKGEAAERPEA 21
|:|:|:|:|:|:|
Db 3 KAEAKAEAKAEAKAE 18

RESULT 6
US-09-218-950-52
; Sequence 52, Application US/09218950
; Patent No. 6284240
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/218,950
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,391
; FILING DATE: 02-AUG-1994
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-218-950-52

Query Match 32.2%; Score 37; DB 3; Length 20;
Best Local Similarity 56.2%; Pred. No. 22;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 6 KTAAKGEAAARPGEA 21
| | | | | | | | | |
DB 3 KAAKAAKAAKAAKAE 18

RESULT 7
US-08-346-849-33
; Sequence 33, Application US/08346849
; Patent No. 5670483
; GENERAL INFORMATION:
; APPLICANT: Zhang, Shuguang
; APPLICANT: Lockshin, Curtis
; APPLICANT: Rich, Alexander
; APPLICANT: Holmes, Todd
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/346,849
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,326
; FILING DATE: 28 DECEMBER 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REFERENCE/DOCKET NUMBER: MIT-6008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-346-849-33

Query Match 29.6%; Score 34; DB 1; Length 16;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 SKTAAKGEAAARPGE 20
: | | | | | | | | | |
DB 1 AKAAKAAKAAKAAKAE 16

RESULT 8
US-08-293-284A-33
; Sequence 33, Application US/08293284A
; Patent No. 5955343
; GENERAL INFORMATION:
; APPLICANT: Holmes, Todd
; APPLICANT: Zhang, Shuguang

; APPLICANT: Rich, Alexander
; APPLICANT: DiPersio, C. Michael
; APPLICANT: Lockshin, Curtis
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/293,284A
; FILING DATE: 22-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,326
; FILING DATE: 28-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: MIT-6008A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-293-284A-33

Query Match 29.6%; Score 34; DB 2; Length 16;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 SKTAAKGEAAARPGE 20
: | | | | | | | | | |
DB 1 AKAAKAAKAAKAAKAE 16

RESULT 9
US-08-898-300-33
; Sequence 33, Application US/08898300
; Patent No. 6548630
; GENERAL INFORMATION:
; APPLICANT: Zhang, Shuguang
; APPLICANT: Lockshin, Curtis
; APPLICANT: Rich, Alexander
; APPLICANT: Holmes, Todd
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/898,300
FILING DATE: 22 JULY 1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/346,849
FILING DATE: 30 NOVEMBER 1994
PRIOR APPLICATION DATA: 07/973,326
APPLICATION NUMBER: 28 DECEMBER 1992
FILING DATE: 28 DECEMBER 1992
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: MIT-6008FB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-898-300-33

Query Match 29.6%; Score 34; DB 4; Length 16;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 SKTAAKGEAAARPE 20
: ||| ||| |||
Db 1 AKAEAKAEAKAEAK 16

RESULT 10

US-08-557-309B-58
Sequence 58, Application US/08557309B
Patent No. 5916572

GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond L.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/557,309B
FILING DATE: 14-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.422
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-557-309B-58

Query Match 29.6%; Score 34; DB 2; Length 21;
Best Local Similarity 47.4%; Pred. No. 67;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 KTAAKGEAAARPEGAAVA 24
: ||| ||| ||| |||
Db 1 KAAAPAKAAAPAKAAAA 19

RESULT 11

US-08-929-414-11
Sequence 11, Application US/08929414
Patent No. 5942403

GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Haughton, Raymond
APPLICANT: Skeiky, Yasir A.W.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION
OF T. CRUZI INFECTION
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/929,414
FILING DATE: 15-SEP-1997

CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.406C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-929-414-11

Query Match 29.6%; Score 34; DB 2; Length 21;
Best Local Similarity 47.4%; Pred. No. 67;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 KTAAKGEAAARPEGAAVA 24
: ||| ||| ||| |||
Db 1 KAAAPAKAAAPAKAAAA 19

RESULT 12

US-08-834-306-56
Sequence 56, Application US/08834306
Patent No. 6054135

GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Lodes, Michael J.

```

; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,306
; FILING DATE: 15-APR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.422C1
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-834-306-56

Query Match 29.6%; Score 34; DB 3; Length 21;
Best Local Similarity 47.4%; Pred. No. 67;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 KTAAKGEAAARPPGEAAVA 24
DB 1 KAAAPAKAAAPAKAAAA 19

RESULT 13
US-08-834-306-58
; Sequence 58, Application US/08834306
; Patent No. 6054135
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,306
; FILING DATE: 15-APR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392

```

```

; REFERENCE/DOCKET NUMBER: 210121.422C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-834-306-58

Query Match 29.6%; Score 34; DB 3; Length 21;
Best Local Similarity 47.4%; Pred. No. 67;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 KTAAKGEAAARPPGEAAVA 24
DB 1 KAAAPAKAAAPAKAAAA 19

RESULT 14
US-08-993-674A-56
; Sequence 56, Application US/08993674A
; Patent No. 6228372
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Smith, John M.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,674A
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.422C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-993-674A-56

Query Match 29.6%; Score 34; DB 3; Length 21;
Best Local Similarity 47.4%; Pred. No. 67;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 KTAAKGEAAARPPGEAAVA 24
DB 1 KAAAPAKAAAPAKAAAA 19

```

```

RESULT 15
US-08-993-674A-58
; Sequence 58, Application US/08993674A
; Patent No. 6228372
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Smith, John M.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,674A
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.422C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-993-674A-58

Query Match 29.6%; Score 34; DB 3; Length 21;
Best Local Similarity 47.4%; Pred.No. 67;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 KTAAGGEAAAEERPGEEAAVA 24
Db 1 KAAAPAPAKAAAPAKAAAA 19

Search completed: October 7, 2004, 17:58:13
Job time : 23 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:47:47 ; Search time 23.5 seconds
(without alignments)
98.238 Million cell updates/sec

Title: US-10-802-644-2

Perfect score: 115

Sequence: 1 GTAPAAEGAGAEVKRASAEAKQAF 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 4740

Minimum DB seq length: 0

Maximum DB seq length: 24

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%	Query Match	Length	ID	Description
1	37	32.2	24	2	S33302	hydrogenase (EC 1.
2	36	31.3	20	2	S38291	30K allergen - vel
3	36	31.3	22	2	D47256	kinetoplast DNA-as
4	32	27.8	23	2	I64839	SP-A1 (alpha, beta
5	30.5	26.5	20	2	S38294	32K allergen - com
6	30	26.1	24	2	A61505	pollen allergen Ph
7	30	26.1	24	2	F45357	Kex2/subtilisin-li
8	30	26.1	24	2	C45357	Kex2/subtilisin-li
9	29	25.2	20	2	S19618	globin - polychaet
10	28	24.3	15	2	PQ0692	photosystem I 18.5
11	28	24.3	16	2	S38292	30K allergen - rye
12	28	24.3	21	2	S32883	DNA gyrase chain A
13	26.5	23.0	22	2	A26376	peptidyl-di-peptida
14	26	22.6	20	2	PH0918	T-cell receptor be
15	26	22.6	15	2	FS0260	56K protein 2402 -
16	26	22.6	17	2	C30221	histone H2A.8 - ch
17	26	22.6	18	2	S45373	translation elonga
18	26	22.6	20	2	PQ0687	photosystem I 14.1
19	25	21.7	16	2	A11488	taurocyamine kinas
20	25	21.7	20	2	S32502	calpain (EC 3.4.22
21	24.5	21.3	19	2	B29501	fibriopeptide A -
22	24	20.9	13	2	S23638	Ig kappa chain J s
23	24	20.9	20	2	S06149	photosystem I chai
24	24	20.9	20	2	S38288	50K allergen - per
25	24	20.9	20	2	PC1150	equinotoxin 1B - s
26	23	20.0	10	2	H60588	sperm-activating p
27	23	20.0	12	1	A43975	locustamytotropin -
28	23	20.0	12	2	PH1481	T-cell receptor be
29	23	20.0	14	2	S23639	Ig kappa chain J s

30	23	20.0	14	2	E81280	probable proteolys
31	23	20.0	15	2	PA0087	cytochrome c2 - fu
32	23	20.0	15	2	A49177	22K protein p1, m
33	23	20.0	16	2	PA0103	L-lactate dehydrog
34	23	20.0	19	2	H27480	hydrogenase (EC 1.
35	23	20.0	20	2	S33787	pancreatic elastas
36	23	20.0	20	2	B44920	2-halobenzoate 1,2
37	23	20.0	20	2	A58903	metalloprotease
38	23	20.0	22	2	S68616	histone H1 - sea u
39	23	20.0	22	2	PC2134	maltose transport
40	23	20.0	23	2	S38991	glycine reductase
41	23	20.0	24	2	S70351	peptidyl-di-peptida
42	22	19.1	12	2	A20907	Ig kappa chain J1
43	22	19.1	13	2	B58533	CD61 homolog - cha
44	22	19.1	13	2	B20907	Ig kappa-1 chain J
45	22	19.1	13	2	B25448	Ig kappa-1 chain,

ALIGNMENTS

RESULT 1

S33302 hydrogenase (EC 1.18.99.1) - Chlamydomonas reinhardtii (fragment)
C:Species: Chlamydomonas reinhardtii

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 30-Sep-1993
C:Accession: S33302

R:Happe, T.; Naber, J.D.

Eur. J. Biochem. 214, 475-481, 1993

A:Title: Isolation, characterization and N-terminal amino acid sequence of hydrogenase fi

A:Reference number: S33302; MUID:93292509; PMID:8513797

A:Accession: S33302

A:Molecule type: protein

A:Residues: 1-24 <HAP>

C:Keywords: chloroplast; oxidoreductase

Query Match 32.2%; Score 37; DB 2; Length 24;

Best Local Similarity 52.9%; Pred. No. 1.1e+02;

Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 APAAEGAGAEVKRASAE 19

DB 2 APAAEAPLSHVQALAE 18

RESULT 2

S38291 30K allergen - velvet grass (fragment)

C:Species: Holcus lanatus (velvet grass)

C:Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 19-May-2000

C:Accession: S38291

R:Petersen, A.; Schramm, G.; Becker, W.M.; Schlaak, M.

Biol. Chem. Hoppe-Seyler 374, 855-861, 1993

A:Title: Comparison of four grass pollen species concerning their allergens of grass gro

A:Reference number: S38288; MUID:94092339; PMID:7505588

A:Accession: S38291

A:Molecule type: protein

A:Residues: 1-20 <PET>

Query Match 31.3%; Score 36; DB 2; Length 20;

Best Local Similarity 70.0%; Pred. No. 1.3e+02;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 APAAEGAGAE 12

DB 9 APAATGAGGD 18

RESULT 3

D47256 kinetoplast DNA-associated protein p15 - Crithidia fasciculata (fragment)

C:Species: Crithidia fasciculata

C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994

C;Accession: D47256
R;Xu, C.; Ray, D.S.
Proc. Natl. Acad. Sci. U.S.A. 90, 1786-1789, 1993
A;Title: Isolation of proteins associated with kinetoplast DNA networks in vivo.
A;Reference number: A47256; MUID:93189582; PMID:8446592
A;Accession: D47256
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-22 <XU1>
A;Note: sequence extracted from NCBI backbone (NCBIP:126909)

Query Match 31.3%; Score 36; DB 2; Length 22;
Best Local Similarity 42.1%; Pred. No. 1.4e+02;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 APAAGAGAEVKRASAEAK 21
||| : : : : |||
Db 4 APAKAAAPKAAKASTPAK 22

RESULT 4
I64839
SP-A1 (alpha, beta, epsilon) and SP-A2 (alpha, beta) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence revision 29-May-1998 #text_change 20-Aug-1999
C;Accession: I64839; I64842; I64841; I64845; I64838
R;McCormick, S.M.; Boggaram, V.; Mendelson, C.R.
Am. J. Physiol. 266, L354-L366, 1994
A;Title: Characterization of mRNA transcripts and organization of human SP-A1 and SP-A2
A;Reference number: I51910; MUID:94234365; PMID:8179012
A;Accession: I64839
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-23 <RES>
A;Cross-references: GB:S69681; NID:g546673; PIDN:AAB30731.1; PID:g546674
A;Accession: I64842
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-23 <RE2>
A;Cross-references: GB:S69686; NID:g546681; PIDN:AAB30735.1; PID:g546682
A;Accession: I64841
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-23 <RE4>
A;Cross-references: GB:S69690; NID:g546687; PIDN:AAB30738.1; PID:g546688
A;Accession: I64838
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-23 <RE5>
A;Cross-references: GB:S69680; NID:g546671; PIDN:AAB30730.1; PID:g546672
C;Genetics:
A;Gene: SP-A2; SP-A1
C;Superfamily: mannose-binding lectin; C-type lectin homology

Query Match 27.8%; Score 32; DB 2; Length 23;
Best Local Similarity 70.0%; Pred. No. 5.3e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 AAGAGAEVK 14
||| |||
Db 14 AAGAAACEVK 23

RESULT 5
S38294
32K allergen - common timothy (fragment)
C;Species: Phleum pratense (common timothy)
C;Date: 19-May-1994 #sequence_revision 27-Feb-1997 #text_change 07-May-1999

C;Accession: S38294
R;Petersen, A.; Schramm, G.; Becker, W.M.; Schlaak, M.
Biol. Chem. Hoppe-Seyler 374, 855-861, 1993
A;Title: Comparison of four grass pollen species concerning their allergens of grass group 1.
A;Reference number: S38288; MUID:94092339; PMID:7505588
A;Accession: S38294
A;Molecule type: protein
A;Residues: 1-20 <PET>

Query Match 26.5%; Score 30.5; DB 2; Length 20;
Best Local Similarity 52.9%; Pred. No. 7.5e+02;
Matches 9; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

QY 1 GTAPAA-EGAGAEVKRA 16
||| ||| : :
Db 4 GYAPATPAAGAEAGKA 20

RESULT 6
A61505
pollen allergen Phl p V - common timothy (fragment)
C;Species: Phleum pratense (common timothy)
C;Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 07-May-1999
C;Accession: A61505
R;Matthiesen, F.; Lowenstein, H.
Clin. Exp. Allergy 21, 297-307, 1991
A;Title: Group V allergens in grass pollens. I. Purification and characterization of the
A;Reference number: A61505; MUID:91322736; PMID:1863892
A;Accession: A61505
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-24 <MAT>
A;Note: 2-Ala, 5-Gly, and 15-Glu were also found
C;Superfamily: grass pollen allergen IX
C;Keywords: hydroxyproline; pollen
F;6,9,12,18,21,24/Modified site: hydroxyproline (Pro) #status experimental

Query Match 26.1%; Score 30; DB 2; Length 24;
Best Local Similarity 85.7%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAAGAG 10
||| |||
Db 9 PAAPGAG 15

RESULT 7
F45357
Kex2/subtilisin-like proprotein convertase PC4-C - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C;Accession: F45357
R;Seidah, N.G.; Day, R.; Hamelin, J.; Gaspar, A.; Collard, M.W.; Chretien, M.
Mol. Endocrinol. 6, 1559-1570, 1992
A;Title: Testicular expression of PC4 in the rat: molecular diversity of a novel germ cell
A;Reference number: A45357; MUID:93078790; PMID:1448111
A;Accession: F45357
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-24 <SE1>
A;Note: sequence extracted from NCBI backbone (NCBIP:118888)

Query Match 26.1%; Score 30; DB 2; Length 24;
Best Local Similarity 31.2%; Pred. No. 1e+03;
Matches 5; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 TAPAAEGAGAEVKRAS 17
::||| : : :
Db 8 SSPASSGGGSTATSS 23

RESULT 8
C45357


```

Kex2/subtilisin-like proprotein convertase PC4-C - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C:Accession: C45357
R:Seidah, N.G.; Day, R.; Hamelin, J.; Gaspar, A.; Collard, M.W.; Chretien, M.
Mol. Endocrinol. 6, 1559-1570, 1992
A:Title: Testicular expression of PC4 in the rat: molecular diversity of a novel germ cell
A:Reference number: A45357; MUID:93078790; PMID:1448111
A:Accession: C45357
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-24 <SEI>
A>Note: sequence extracted from NCBI backbone (NCBIP:118879)

Query Match 26.1%; Score 30; DB 2; Length 24;
Best Local Similarity 31.2%; Pred. No. 1e+03;
Matches 5; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 TAPAAEGAGAEVKRAS 17
   :||: ||: |
Db 8 SSPASSGGGGSTATHSS 23

RESULT 9
S19618
globin - polychaete (Eudistylia vancouveri) (fragment)
N:Alternate names: chlorocruorin
C:Species: Eudistylia vancouveri
C>Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 05-Dec-1998
C:Accession: S19618
R:Qabar, A.N.; Stern, M.S.; Walz, D.A.; Chiu, J.T.; Timkovich, R.; Wall, J.S.; Kapp, O.H.
J. Mol. Biol. 222, 1109-1129, 1991
A:Title: Hierarchy of globin complexes. The quaternary structure of the extracellular chain
A:Reference number: S19532; MUID:92106333; PMID:1762147
A:Accession: S19618
A:Molecule type: protein
A:Residues: 1-20 <QAB>
A:Experimental source: plume
C:Complex: dodecamer, each consisting of a trimer of tetramers of globin chains; dodecamer
C:Keywords: calcium; dodecamer; heme; homotetramer; oxygen carrier

Query Match 25.2%; Score 29; DB 2; Length 20;
Best Local Similarity 43.8%; Pred. No. 1.2e+03;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 7 EGAGAEVVKRAAEAKQ 22
   | : ||||| |
Db 1 ELSSSEVKRIDANGKK 16

RESULT 10
PQ0692
photosystem I 18.5K D2 chain - tobacco (Nicotiana glauca) (fragment)
C:Species: Nicotiana glauca
C>Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 17-Mar-1999
C:Accession: PQ0692
R:Obokata, J.; Mikami, K.; Hayashida, N.; Nakamura, M.; Sugiyama, M.
Plant Physiol. 102, 1259-1267, 1993
A:Title: Molecular heterogeneity of photosystem I. psaD, psaE, psaF, psaH and psaL are a
A:Reference number: PQ0667; MUID:94105345; PMID:8278548
A:Accession: PQ0692
A:Molecule type: protein
A:Residues: 1-15 <QBO>
C:Superfamily: photosystem I chain II
C:Keywords: chloroplast; photosynthesis; photosystem I; thylakoid

Query Match 24.3%; Score 28; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 7; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 AAEAGAGAEVKRASA 18
   | | | | | |
Db 1 AEEAATAATKEAEA 14

```

RESULT 11

```

S38292
30K allergen - rye (fragment)
C:Species: Secale cereale (rye)
C>Date: 19-May-1994 #sequence_revision 27-Feb-1997 #text_change 07-May-1999
C:Accession: S38292
R:Petersen, A.; Schramm, G.; Becker, W.M.; Schlaak, M.
Biol. Chem. Hoppe-Seyler 374, 855-861, 1993
A:Title: Comparison of four grass pollen species concerning their allergens of grass group
A:Reference number: S38288; MUID:94092339; PMID:7505588
A:Accession: S38292
A:Molecule type: protein
A:Residues: 1-16 <PEI>

```

```

Query Match 24.3%; Score 28; DB 2; Length 16;
Best Local Similarity 75.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 GTAPAAEG 8
   | | | | |
Db 4 GYAPAAPG 11

```

RESULT 12

```

S32883
DNA gyrase chain A - Streptomyces spheroides (fragment)
C:Species: Streptomyces spheroides
C>Date: 08-Dec-1993 #sequence_revision 26-May-1995 #text_change 22-Oct-1999
C:Accession: S32883; S29684
R:Thiara, A.S.; Cundliffe, E.
Mol. Microbiol. 8, 495-506, 1993
A:Title: Expression and analysis of two gyrB genes from the novobiocin producer, Streptomyces
A:Reference number: S32881; MUID:93316846; PMID:8392138
A:Accession: S32883
A:Molecule type: DNA
A:Residues: 1-21 <THI>
A:Cross-references: EMBL:217305; NID:947535; PIDN:CAA78953.1; PID:947537

```

```

Query Match 24.3%; Score 28; DB 2; Length 21;
Best Local Similarity 71.4%; Pred. No. 1.7e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 4 PAAEGAG 10
   | | | | |
Db 14 PAVEGVG 20

```

RESULT 13

```

A26376
peptidyl-dipeptidase A (EC 3.4.15.1) - bovine (fragment)
N:Alternate names: ACE; angiotensin-converting enzyme; dipeptidyl carboxypeptidase I; per
C:Species: Bos primigenius taurus (cattle)
C>Date: 21-May-1988 #sequence_revision 05-Apr-1995 #text_change 30-Jun-1995
C:Accession: A26376; B61477
R:St. Clair, D.K.; Presper, K.A.; Smith, P.L.; Stump, D.C.; Heath, E.C.
Biochem. Biophys. Res. Commun. 141, 968-972, 1986
A:Title: Bovine angiotensin-converting enzyme: amino terminal sequence analysis and preli
A:Reference number: A26376; MUID:87128067; PMID:3028395
A:Accession: A26376
A:Molecule type: protein
A:Residues: 1-22 <STC>
A:Experimental source: lung
R:Bernstein, K.B.; Martin, B.M.; Striker, L.; Striker, G.
Kidney Int. 33, 652-655, 1988
A:Title: Partial protein sequence of mouse and bovine kidney angiotensin converting enzy
A:Reference number: A61477; MUID:88215372; PMID:2835538
A:Accession: B61477
A:Molecule type: protein
A:Residues: 1-22 <BER>
A:Experimental source: lung
C:Superfamily: mammalian peptidyl-dipeptidase A

```

Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:36:42 ; Search time 14 Seconds
(without alignments)
89.263 Million cell updates/sec

Title: US-10-802-644-2

Perfect score: 115

Sequence: 1 GTAPAAEGAGAEVVKRASAEAKQAF 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 1550

Minimum DB seq length: 0

Maximum DB seq length: 24

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	26.1	19	1	ATPB CANFA
2	30	26.1	21	1	PIL3_ECOLI
3	28	24.3	21	1	GYRA STRSH
4	26.5	23.0	22	1	ACE BOVIN
5	25	21.7	16	1	KTRC AREMA
6	25	21.7	20	1	DER6 DERPPT
7	25	21.7	24	1	AMBA BACTR
8	24.5	21.3	19	1	FIBA BISBO
9	24.5	21.3	19	1	FIBA CEREL
10	24.5	21.3	19	1	FIBA CERNI
11	24	20.9	15	1	R13A SPIOL
12	24	20.9	15	1	UN04 PINPS
13	23	20.0	12	1	H2AX_ONCMY
14	23	20.0	12	1	LM71_LOCFI
15	23	20.0	17	1	UP34_UPEMJ
16	23	20.0	19	1	PHSL_DESEN
17	23	20.0	20	1	ELAS_GADMO
18	22	19.1	21	1	TL19_ARATH
19	22	19.1	23	1	VG22_BPT2
20	22	19.1	23	1	VG22_BPT6
21	21.5	18.7	19	1	FIBA MUNMU
22	21	18.3	8	1	WP1_PERAT
23	21	18.3	12	1	PORD_METTM
24	21	18.3	12	1	TAP0_TREME
25	21	18.3	14	1	SAP2_ARBPB
26	21	18.3	14	1	UC04_MAIZE
27	21	18.3	15	1	FIBA SYNCA
28	21	18.3	15	1	RS20_BACST
29	21	18.3	16	1	CERR_RAT
30	21	18.3	16	1	FIBA_CERSI
31	21	18.3	16	1	FIBA_FELCA
32	21	18.3	16	1	FIBA_HYLLA
33	21	18.3	16	1	FIBA_NACFU

34	21	18.3	16	1	FIBA_ODOHE
35	21	18.3	16	1	FIBA_TAPTE
36	21	18.3	17	1	FIBA_PIG
37	21	18.3	18	1	FIBA_CAMDR
38	21	18.3	18	1	FIBA_LAMGL
39	21	18.3	18	1	YMDB_CHLAU
40	21	18.3	19	1	FIBA_BUBBU
41	21	18.3	19	1	FIBA_SHEEP
42	21	18.3	20	1	FIBB_ELEMA
43	21	18.3	21	1	PEDB_HYDAT
44	21	18.3	21	1	PSAL_SYNPG
45	21	18.3	21	1	SOD2_PICAB

ALIGNMENTS

RESULT 1
ATPB CANFA STANDARD; PRT; 19 AA.
AC P99504;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP synthase beta chain, mitochondrial (EC 3.6.3.14) (Fragment).
GN ATP5B.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE.
RC TISSUE=Heart;
RX MEDLINE=98163340; PubMed=9504812;
RA Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of dog heart proteins.";
RL Electrophoresis 18:2795-2802(1997).
CC -!- FUNCTION: Produces ATP from ADP in the presence of a proton gradient across the membrane. The beta chain is the catalytic subunit.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate + H(+) (Out).
CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic core - and CF(0) - the membrane proton channel. CF(1) has five subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0) has three main subunits: a, b and c.
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -!- SIMILARITY: Belongs to the ATPase alpha/beta chains family.
CC HSC-2DPAGE; P99504; DOG.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; PARTIAL.
KW ATP synthesis; CF(1); Hydrogen ion transport; Hydrolase; ATP-binding; Mitochondrion.
FT UNSURE 8 8
FT NON_TER 17 19
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 1871 MW; BB9C163FDC60BB42 CRC64;
Query Match 26.1%; Score 30; DB 1; Length 19;
Best Local Similarity 42.9%; Pred. No. 4.2e+02;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 2 TAPAAEGAGAEVVKR 15
Db 4 TSPSPKGAAXXXR 17
RESULT 2
PIL3_ECOLI STANDARD; PRT; 21 AA.
ID PIL3_ECOLI
AC P13948;
DT 01-JAN-1990 (Rel. 13, Created)

```
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Fimbrial protein precursor (Pilin) (Fragment).
GN TRAA.
OS Escherichia coli.
OG Plasmid ColB4-K38.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87008371; PubMed=35311163;
RA Finlay B.B., Frost L.S., Parachych W.;
RT "Origin of transfer of IncF plasmids and nucleotide sequences of the
RT type II oriT, traM, and traI alleles from ColB4-K38 and the type IV
RT traI allele from R100-1.";
RL J. Bacteriol. 168:132-139(1986).
CC -!- FUNCTION: PROPILIN IS THE PRECURSOR OF THE SEXPIILUS SUBUNIT. THE
CC SEXPIILUS ARE FILAMENTOUS SURFACE APPENDAGES REQUIRED FOR CELL-TO-
CC CELL CONTACT DURING BACTERIAL CONJUGATION.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M15135; AAB04666.1; -.
CC InterPro; IPR008873; TrAa.
CC Pfam; PF05513; TrAa; 1.
CC PROSEQ 1 >21
FT PROPEP 21 21
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2198 MW; 076C0D0C5E9D14EA CRC64;

Query Match 26.1%; Score 30; DB 1; Length 21;
Best Local Similarity 46.2%; Pred. No. 4.6e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 AAEGAGAEVKRAS 17
Db :|||:|:|:|
6 SVQGASAPVKKS 18

RESULT 3
GYRA_STRGH STANDARD; PRT; 21 AA.
AC P50071;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE DNA gyrase subunit A (EC 5.99.1.3) (Fragment).
GN GYRA.
OS Streptomyces sphearoides.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=195949;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NCIB 11891;
RX MEDLINE=93316846; PubMed=8392138;
RA Thiara A.S., Cundliffe E.;
RT "Expression and analysis of two gyrB genes from the novobiocin
RT producer, Streptomyces sphearoides.";
RL Mol. Microbiol. 8:495-506(1993).
CC -!- FUNCTION: DNA gyrase negatively supercoils closed circular double-
CC stranded DNA in an ATP-dependent manner and also catalyzes the
CC interconversion of other topological isomers of double-stranded
CC DNA rings, including catenanes and knotted rings.
CC -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
```

```
CC of double-stranded DNA.
CC -!- SUBUNIT: Made up of two chains. The A chain is responsible for DNA
CC breakage and rejoining; the B chain catalyzes ATP hydrolysis. The
CC enzyme forms an A2B2 tetramer.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z17305; CAA78953.1; -.
CC PIR: S32883; S32883.
KW Topoisomerase; isomerase; DNA-binding.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2203 MW; 5E2F9DEA8DB83697 CRC64;

Query Match 24.3%; Score 28; DB 1; Length 21;
Best Local Similarity 71.4%; Pred. No. 8.7e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PAAGGAG 10
Db |||||
14 PAVEGVG 20

RESULT 4
ACE_BOVIN STANDARD; PRT; 22 AA.
ID ACE_BOVIN
AC P12820;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Angiotensin-converting enzyme, somatic (EC 3.4.15.1) (ACE) (Dipeptidyl
DE carboxypeptidase I) (Kininase II) (Fragment).
GN ACE OR DCPI.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=lung;
RX MEDLINE=87128067; PubMed=3028395;
RA St Clair D.K., Presper K.A., Smith P.L., Stump D.C., Heath E.C.;
RT "Bovine angiotensin-converting enzyme: amino-terminal sequence
RT analysis and preliminary characterization of a hybridization-selected
RT primary translation product.";
RL Biochem. Biophys. Res. Commun. 141:968-972(1986).
CC -!- FUNCTION: Converts angiotensin I to angiotensin II by release of
CC the terminal His-Ieu, this results in an increase of the
CC vasoconstrictor activity of angiotensin.
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
CC oligopeptide-|-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither
CC Asp nor Glu. Converts angiotensin I to angiotensin II.
CC -!- COFACTOR: Binds 2 zinc ions per subunit (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family M2.
CC PIR: A26376; A26376.
CC MEROPS; M02.001; -.
CC InterPro; IPR006025; Pept M Zn Bs.
CC PROSITE; PS00142; ZINC_PROTEASE; PARTIAL.
KW Hydrolase; Metalloprotease; Carboxypeptidase; Zinc.
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2271 MW; 77B0AC6A91A3893F CRC64;

Query Match 23.0%; Score 26.5; DB 1; Length 22;
Best Local Similarity 53.8%; Pred. No. 1.4e+03;
Matches 7; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 1 GTAPAAEGAGAEV 13
```

```

Db          9 GNFADE-AGAQI 20
          ||| ||| |||
RESULT 5
KTRC AREMA          STANDARD;          PRT;          16 AA.
AC P11917;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Taurocyamine kinase (EC 2.7.3.4) (Fragment).
OS Arenicola marina (Lugworm) (Rock worm).
OC Eukaryota; Metazoa; Annelida; Polychaeta; Scolecida; Capitellida;
OC Arenicolidae; Arenicola.
OX NCBI_TaxID=6344;
RN [1]
RP SEQUENCE.
RX MEDLINE=75184095; PubMed=166684;
RA Brevet A., Zeitoun Y., Pradel L.A.;
RT "Comparative structural studies of the active site of ATP: guanidine
RT phosphotransferases. The essential cysteine tryptic peptide of
RT taurocyamine kinase from Arenicola marina.";
RL Biochim. Biophys. Acta 393:1-9(1975).
CC -!- CATALYTIC ACTIVITY: ATP + taurocyamine = ADP + N-
CC phosphotaurocyamine.
CC -!- SIMILARITY: Belongs to the ATP:guanido phosphotransferase family.
DR PIR; A11488; A11488.
DR InterPro; IPR000749; ATP-gua Pfams.
DR PROSITE; PS00112; GUANIDO_KINASE; 1.
KW Kinase; Transferase.
FT NON_TER 1 1
FT ACT_SITE 7 7
FT NON_TER 16 16
SQ SEQUENCE 16 AA; 1636 MW; CP057326D427B94 CRC64;

Query Match 21.7%; Score 25; DB 1; Length 16;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTAPAAEGAG 10
Db ||| ||| |||
5 GTCPTNIGTG 14

RESULT 6
DER6_DERPT
ID DER6_DERPT STANDARD; PRT; 20 AA.
AC P49277;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Mite allergen Der p 6 (EC 3.4.21.-) (Der p VI) (DP6) (Fragment).
GN DERP6.
OS Dermatophagoides pteronyssinus (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcoptiformes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.
OX NCBI_TaxID=6956;
RN [1]
RP SEQUENCE.
RX MEDLINE=93327207; PubMed=8334537;
RA Yasueda H., Mita H., Akiyama K., Shida T., Ando T., Sugiyama S.,
RA Yamakawa H.;
RT "Allergens from Dermatophagoides mites with chymotryptic activity.";
RL Clin. Exp. Allergy 23:384-390(1993).
CC -!- CATALYTIC ACTIVITY: Specificity similar to chymotrypsin.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR MEROPS; S01.187; -.
DR InterPro; IPR001254; Peptidase_S1.
DR PROSITE; PS50240; TRYPSIN_DOM; PARTIAL.

```

```

DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
KW Hydrolase; Serine protease; Allergen.
FT UNSURE 1 1
FT UNSURE 6 6
FT UNSURE 10 10
FT UNSURE 11 11
FT UNSURE 17 17
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2082 MW; 26D5F39E06639983 CRC64;

Query Match 21.7%; Score 25; DB 1; Length 20;
Best Local Similarity 71.4%; Pred. No. 2.1e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTAPAAE 7
Db | ||| |||
3 GXQPAAE 9

RESULT 7
AMAA_BACTR
ID AMAA_BACTR STANDARD; PRT; 24 AA.
AC P37356;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE N-acyl-L-amino acid amidohydrolase (EC 3.5.1.14) (L-aminoacylase)
DE (Fragment).
OS Bacillus thermoglucosidasius (Geobacillus thermoglucosidasius).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
OX NCBI_TaxID=1426;
RN [1]
RP SEQUENCE, AND CHARACTERIZATION.
RA Cho H.-Y., Tanizawa K., Tanaka H., Soda K.;
RT "Thermostable aminoacylase from Bacillus thermoglucosidius.
RT Purification and characterization.";
RL Agric. Biol. Chem. 51:2793-2800(1987).
CC -!- CATALYTIC ACTIVITY: An N-acyl-L-amino acid + H(2)O = a fatty acid
CC anion + an L-amino acid.
CC -!- COFACTOR: Activated by cobalt (By similarity).
CC -!- SUBUNIT: Homotetramer.
CC -!- SIMILARITY: Belongs to peptidase family M40.
KW Hydrolase; Cobalt.
FT NON_TER 24 24
SQ SEQUENCE 24 AA; 2846 MW; BC954E4D2B0EC64A CRC64;

Query Match 21.7%; Score 25; DB 1; Length 24;
Best Local Similarity 45.5%; Pred. No. 2.5e+03;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 12 EVKRASAEAKQ 22
Db |::||| |||
5 EIKRLVDEVKE 15

RESULT 8
FIBA_BISO
ID FIBA_BISO STANDARD; PRT; 19 AA.
AC P14441;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibrinogen alpha chain (Contains: Fibrinopeptide A) (Fragment).
GN FGA.
OS Bison bonasus (European bison).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bison.
OX NCBI_TaxID=9902;
RN [1]
RP SEQUENCE.
RA Blomback B., Blomback M., Grondahl N.J.;

```

```

RT "Studies on fibrinopeptides from mammals.";
RL Acta Chem. Scand. 19:1789-1791(1965).
CC -!- FUNCTION: Fibrinogen has a double function: yielding monomers that
CC polymerize into fibrin and acting as a cofactor in platelet
CC aggregation.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
CC which cleaves fibrinopeptides A and B from alpha and beta chains,
CC and thus exposes the N-terminal polymerization sites responsible
CC for the formation of the soft clot.
CC Blood coagulation; Plasma.
KW PEPTIDE 1 19 FIBRINOPEPTIDE A.
FT NON TER 19 19
SQ SEQUENCE 19 AA; 1836 MW; 9BA55A0F473B59C5 CRC64;

Query Match 21.3%; Score 24.5; DB 1; Length 19;
Best Local Similarity 46.7%; Pred. NO. 2.3e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

Qy 1 GTAPA-----AEGAG 10
Db 3 GSDPASGDFLAEGGG 17

RESULT 9
FIBA CEREL
ID FIBA CEREL STANDARD; PRT; 19 AA.
AC P1446;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).
GN FGA.
OS Cervus elaphus (Red deer), and
OS Cervus elaphus nelsoni (American elk).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
CC Cervidae; Cervinae; Cervus.
OX NCBI_TaxID=9860, 9864;
RN [1]
RP SEQUENCE.
RC SPECIES=C.elaphus;
RA Moss G.A., Doolittle R.F.;
RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";
RL Arch. Biochem. Biophys. 122:674-684(1967).
CC -!- FUNCTION: Fibrinogen has a double function: yielding monomers that
CC polymerize into fibrin and acting as a cofactor in platelet
CC aggregation.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
CC which cleaves fibrinopeptides A and B from alpha and beta chains,
CC and thus exposes the N-terminal polymerization sites responsible
CC for the formation of the soft clot.
KW Blood coagulation; Plasma.
FT PEPTIDE 1 19 FIBRINOPEPTIDE A.
FT NON TER 19 19
SQ SEQUENCE 19 AA; 1808 MW; 9BA54C26873B59C5 CRC64;

Query Match 21.3%; Score 24.5; DB 1; Length 19;
Best Local Similarity 46.7%; Pred. NO. 2.3e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

Qy 1 GTAPA-----AEGAG 10
Db 3 GSDPASGDFLAEGGG 17

```

```

RESULT 10
FIBA CERNI
ID FIBA CERNI STANDARD; PRT; 19 AA.
AC P1447;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).
GN FGA.
OS Cervus nippon (Sika deer).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
CC Cervidae; Cervinae; Cervus.
OX NCBI_TaxID=9863;
RN [1]
RP SEQUENCE.
RA Blomback B., Blomback M., Grondahl N.J., Holmberg E.;
RT "Structure of fibrinopeptides-its relation to enzyme specificity and
RT phylogeny and classification of species.";
RL Ark. Kemi 25:411-428(1966).
CC -!- FUNCTION: Fibrinogen has a double function: yielding monomers that
CC polymerize into fibrin and acting as a cofactor in platelet
CC aggregation.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
CC which cleaves fibrinopeptides A and B from alpha and beta chains,
CC and thus exposes the N-terminal polymerization sites responsible
CC for the formation of the soft clot.
KW Blood coagulation; Plasma.
FT NON TER 19 19
SQ SEQUENCE 19 AA; 1822 MW; 9BA40926873B59C5 CRC64;

Query Match 21.3%; Score 24.5; DB 1; Length 19;
Best Local Similarity 46.7%; Pred. NO. 2.3e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

Qy 1 GTAPA-----AEGAG 10
Db 3 GSDPASSEFLAEGGG 17

```

```

RESULT 11
R13A SPIOL
ID R13A SPIOL STANDARD; PRT; 15 AA.
AC P82454;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 60S ribosomal protein L13a (Fragment).
GN RPL13A.
OS Spinacia oleracea (Spinach).
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
CC Caryophyllales; Amaranthaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RC STEAIN=cv. Alvaro; TISSUE=Leaf;
RA Yamaguchi K., Subramanian A.R.;
RT "N-terminal sequence of spinach cytosolic 60S ribosomal protein
RT L13a.";
RL Submitted (APR-2000) to Swiss-Prot.
CC -!- SIMILARITY: Belongs to the L13P family of ribosomal proteins.
DR InterPro; IPR005822; Ribosomal_L13.
DR PROSITE; PS00783; RIBOSOMAL_L13; PARTIAL.
KW Ribosomal protein.
FT NON TER 15 15
SQ SEQUENCE 15 AA; 1489 MW; C7B9C80F5A099EB3 CRC64;

Query Match 20.9%; Score 24; DB 1; Length 15;

```


[1]
 RN SEQUENCE, AND MASS SPECTROMETRY.
 RP TISSUE=Skin secretion;
 RA Bradford A.M.; Bowie J.H., Tyler M.J., Wallace J.C.;
 RT "New antibiotic uperin peptides from the dorsal glands of the
 RT Australian toadlet Uperoleia mjobergii.";
 RL Aust. J. Chem. 49:1325-1331(1996).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
 CC -!- MASS SPECTROMETRY: MW=1735; METHOD=FAB.
 KW Amphibian defense peptide; Amidation.
 FT MOD RES 17 17 AMIDATION.
 SQ SEQUENCE 17 AA; 1737 MW; 6F61E4834375DE1B CRC64;

Query Match 20.0%; Score 23; DB 1; Length 17;
 Best Local Similarity 35.7%; Pred. No. 3.4e+03;
 Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 8 GAGAEVKRASAFAK 21
 | | : : | |
 Db 1 GVGDLIRKAAIAIK 14

Search completed: October 7, 2004, 17:54:03
 Job time : 15 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:37:47 ; Search time 72 Seconds
(without alignments)
105.173 Million cell updates/sec

Title: US-10-802-644-2

Perfect score: 115

Sequence: 1 GTAPAEAGAGAEVKRASAEAKQAF 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 segs, 315518202 residues

Total number of hits satisfying chosen parameters: 10229

Minimum DB seq length: 0

Maximum DB seq length: 24

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPTRMBL 25:*

2: sp_archaea:*

3: sp_bacteria:*

4: sp_fungi:*

5: sp_human:*

6: sp_invertebrate:*

7: sp_mammal:*

8: sp_mmc:*

9: sp_organelle:*

10: sp_phage:*

11: sp_plant:*

12: sp_rodent:*

13: sp_virus:*

14: sp_vertebrate:*

15: sp_unclassified:*

16: sp_virus:*

17: sp_bacteriaph:*

18: sp_archaea:*

19: sp_bacteria:*

20: sp_fungi:*

21: sp_human:*

22: sp_invertebrate:*

23: sp_mammal:*

24: sp_mmc:*

25: sp_organelle:*

26: sp_phage:*

27: sp_plant:*

28: sp_rodent:*

29: sp_virus:*

30: sp_vertebrate:*

31: sp_unclassified:*

32: sp_virus:*

33: sp_bacteriaph:*

34: sp_archaea:*

35: sp_bacteria:*

36: sp_fungi:*

37: sp_human:*

38: sp_invertebrate:*

39: sp_mammal:*

40: sp_mmc:*

41: sp_organelle:*

42: sp_phage:*

43: sp_plant:*

44: sp_rodent:*

45: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	37	32.2	24	10 Q9S8S0	Q9S8S0 chlamydomon
2	36	31.3	22	5 Q9TWU8	Q9TWU8 crithidia f
3	32	27.8	23	4 P78489	P78489 homo sapien
4	32	27.8	23	4 Q9UCN3	Q9UCN3 homo sapien
5	31	27.0	20	4 Q15988	Q15988 homo sapien
6	31	27.0	23	4 Q43519	Q43519 homo sapien
7	31	27.0	24	2 P81151	P81151 desulfovibr
8	30.5	26.5	20	10 Q9S8X9	Q9S8X9 phleum prat
9	30	26.1	15	10 Q9SQI6	Q9SQI6 oryza sativ
10	30	26.1	17	11 Q8K407	Q8K407 rattus norv
11	30	26.1	24	12 Q83877	Q83877 norwalk vir
12	29	25.2	18	11 O55167	O55167 rattus norv
13	29	25.2	22	8 Q9T217	Q9T217 nicotiana s
14	29	25.2	22	10 Q48861	Q48861 oryza sativ
15	28	24.3	15	8 Q9T210	Q9T210 nicotiana t
16	28	24.3	16	11 Q9QUY8	Q9QUY8 cricetulus

ALIGNMENTS

RESULT 1

Q9S8S0 PRELIMINARY; PRT; 24 AA.
AC Q9S8S0
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hydrogenase (EC 1.12.-.-) (Fragment).
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE.
RX MEDLINE=93292509; PubMed=8513797;
RA Happe T., Naber J.D.;
RT "Isolation, characterization and N-terminal amino acid sequence of
RT hydrogenase from the green alga Chlamydomonas reinhardtii.";
RL Eur. J. Biochem. 214:475-481(1993).
DR PIR; S33302; S33302.
SQ SEQUENCE 24 AA; 2427 MW; A582111FEALB69EC CRC64;

Query Match 32.2%; Score 37; DB 10; Length 24;
Best Local Similarity 52.9%; Pred. No. 3.5e+02;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 APAAEGAGAEVKRASAE 19
||||| : : : : :
Db 2 APAEAPLHVQQAALAE 18

RESULT 2

Q9TWU8 PRELIMINARY; PRT; 22 AA.
ID Q9TWU8
AC Q9TWU8
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Kinetoplast DNA-associated protein (Fragment).

OS Crithidia fasciculata.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Crithidia.
OX NCBI_TaxID=5656;
RN [1]
RP SEQUENCE.
RX MEDLINE=93189582; PubMed=8446592;
RA Xu C., Ray D.S.;
RT "Isolation of proteins associated with kinetoplast DNA networks in vivo";
RL Proc. Natl. Acad. Sci. U.S.A. 90:1786-1789(1993).
DR PIR; D47256;
SQ SEQUENCE 22 AA; 1991 MW; 5B1C4941E08F9B2 CRC64;

Query Match 31.3%; Score 36; DB 5; Length 22;
Best Local Similarity 42.1%; Pred. No. 4.4e+02;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 APAEGAGAEVVKRAAEAK 21
||| : : : : :
DB 4 APAKKAAPKAASAPAK 22
||| : : : : :

RESULT 3
P78489 PRELIMINARY; PRT; 23 AA.
AC P78489;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE SP-A1 epsilon (Fragment).
GN SP-A1 OR SP-A2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94234365; PubMed=8179012;
RA McCormick S.M., Boggs V., Mendelson C.R.;
RT "Characterization of mRNA transcripts and organization of human SP-A1 and SP-A2 genes";
RL Am. J. Physiol. 266:L354-L366(1994).
DR EMBL; S69690; AAB30738.1; -
DR EMBL; S69680; AAB30730.1; -
DR EMBL; S69681; AAB30731.1; -
DR EMBL; S69685; AAB30734.1; -
DR EMBL; S69686; AAB30735.1; -
DR PIR; I64839; I64839.
FT NON TER 23
SQ SEQUENCE 23 AA; 2419 MW; 8C4FE026553AAD2C CRC64;

Query Match 27.8%; Score 32; DB 4; Length 23;
Best Local Similarity 70.0%; Pred. No. 1.5e+03;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 ABEGAGAEVK 14
||| ||| |||
DB 14 ABEGAGAEVK 23
||| ||| |||

RESULT 4
Q9UCN3 PRELIMINARY; PRT; 23 AA.
AC Q9UCN3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Interleukin-1 beta converting enzyme (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE.
RX MEDLINE=9337439; PubMed=1321594;
RA Kronheim S.R., Mumma A., Greenstreet T., Glackin P.J., Van Ness K.,
RA March C.J., Black R.A.;
RT "Purification of interleukin-1 beta converting enzyme, the protease that cleaves the interleukin-1 beta precursor";
RL Arch. Biochem. Biophys. 296:698-703(1992).
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000873; AMP-bind.
DR PROSITE; PS00455; AMP_BINDING; 1.
SQ SEQUENCE 23 AA; 2322 MW; 63C2E59F1F3102D6 CRC64;

Query Match 27.8%; Score 32; DB 4; Length 23;
Best Local Similarity 36.8%; Pred. No. 1.6e+03;
Matches 7; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 4 PAEGAGAEVVKRAAEAKQ 22
||| : : : : :
DB 5 PTSSGSEGNVKLXLEXAQ 23
||| : : : : :

RESULT 5
Q15988 PRELIMINARY; PRT; 20 AA.
AC Q15988;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE HLX1 protein (fragment).
GN HLX1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93194183; PubMed=8095486;
RA Nishimura D.Y., Purchio A.F., Murray J.C.;
RT "Linkage localization of TGFB2 and the human homeobox gene HLX1 to chromosome 1q";
RL Genomics 15:357-364(1993).
DR EMBL; S56767; AAD13883.1; -
FT NON TER 1
SQ SEQUENCE 20 AA; 2012 MW; 6BB655F09B5B5AE4 CRC64;

Query Match 27.0%; Score 31; DB 4; Length 20;
Best Local Similarity 75.0%; Pred. No. 1.9e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTAPAAEG 8
||| |||
DB 4 GGNPAADG 11
||| |||

RESULT 6
O43519 PRELIMINARY; PRT; 23 AA.
AC O43519;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE RET proto-oncogene (fragment).
GN RET.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98086192; PubMed=9426223;
RA Patrone G., Puliti A., Boccia R., Ravazzolo R., Romeo G.;
RT "Sequence and characterisation of the RET proto-oncogene 5' flanking

RT region: analysis of retinoic acid responsiveness at the
 RT transcriptional level."
 RL FEBS Lett. 419:76-82(1997).
 [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=99054655; PubMed=9840920;
 RA Munnes M., Patrone G., Schmitz B., Romeo G., Doerfler W.;
 RT "A 5'-CG-3'-rich region in the promoter of the transcriptionally
 RT frequently silenced RET protooncogene lacks methylated cytidine
 RT residues."
 RL Oncogene 17:2573-2583(1998).
 DR EMBL AF032124; AAB97168.1; -.
 FT NON_TER 1 1
 FT NON_TER 23 23
 SQ SEQUENCE 23 AA; 2129 MW; 1DBAF1BFBOE4C66D CRC64;

Query Match 27.0%; Score 31; DB 4; Length 23;
 Best Local Similarity 46.7%; Pred. No. 2.2e+03;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 8 GAGAEVKRASAEAKQ 22
 |||:|:|:|:|:
 Db 9 GAASAVAAAAARQ 23

RESULT 7
 P81151
 ID P81151 PRELIMINARY; PRT; 24 AA.
 AC P81151;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Cytochrome C3 (Fragment).
 OS Desulfovibrio vulgaris.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
 OC Desulfovibrionaceae; Desulfovibrio.
 OX NCBI_TaxID=881;
 RN [1]
 RP SEQUENCE.

RX MEDLINE=93272123; PubMed=8388770;
 RA Kwok D.Y., Vedvick T.S., McCue A.F., Gevertz D.;
 RT "Rapid comparison of the cytochrome c3 gene from nine strains of
 RT Desulfovibrio vulgaris using polymerase chain reaction
 RT amplification."
 RL Can. J. Microbiol. 39:403-411(1993).
 CC -!- FUNCTION: PARTICIPATES IN SULFATE RESPIRATION COUPLED WITH
 CC PHOSPHORYLATION BY TRANSFERRING ELECTRONS FROM THE ENZYME
 CC DEHYDROGENASE TO FERREDOXIN.
 CC -!- PTM: BINDS FOUR NONPARALLEL HEME GROUPS PER MOLECULE.
 CC -!- SIMILARITY: TO OTHER C3-TYPE CYTOCHROMES.
 DR GO; GO:0009061; P:anaerobic respiration; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 KW Electron transport; Sulfate respiration; Heme.
 FT NON_TER 1 1
 FT NON_TER 24 24
 SQ SEQUENCE 24 AA; 2355 MW; B4F438AFF102472 CRC64;

Query Match 27.0%; Score 31; DB 2; Length 24;
 Best Local Similarity 42.9%; Pred. No. 2.3e+03;
 Matches 9; Conservative 3; Mismatches 5; Indels 4; Gaps 1;

QY 1 GTAPAAEGAGAEVKRASAEAK 21
 |||:|:|:|:|:
 Db 1 GNAPAAD----MVLKAPGDAK 17

RESULT 8
 Q9S8X9
 ID Q9S8X9 PRELIMINARY; PRT; 20 AA.
 AC Q9S8X9;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Allergen PHLP V (Fragment).
 OS Phleum pratense (Common timothy).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Aveneae; Phleum.
 OX NCBI_TaxID=15957;
 RN [1]
 RP SEQUENCE
 RX MEDLINE=92353723; PubMed=1643437;
 RA Petersen A., Becker W.M., Schlaak M.;
 RT "Characterization of isoforms of the major allergen Phl p V by two-
 RT dimensional immunoblotting and microsequencing."
 RL Int. Arch. Allergy Immunol. 98:105-109(1992).
 DR PIR: S38294; S38294.
 SQ SEQUENCE 20 AA; 1731 MW; F0DD5E9B482D9836 CRC64;

Query Match 26.5%; Score 30.5; DB 10; Length 20;
 Best Local Similarity 52.9%; Pred. No. 2.3e+03;
 Matches 9; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

QY 1 GTAPAA-EGAGAEVKRA 16
 |||:|:|:|:|:
 Db 4 GYAPATPAAAGAEAGKA 20

RESULT 9
 Q9SQI6
 ID Q9SQI6 PRELIMINARY; PRT; 15 AA.
 AC Q9SQI6;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Translation elongation factor (fragment).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Li Z.-Y., Chen S.-Y.;
 RT "Inducible expression of translation elongation factor 1A gene in rice
 RT seedlings in response to environmental stresses."
 RL Acta Bot. Sin. 41:800-806(1999).
 DR EMBL AF067195; AAC79991.1; -.
 DR Gramene; Q9SQI6; -.
 DR GO; GO:0003746; F:translation elongation factor activity; IEA.
 KW Elongation factor.
 FT NON_TER 1 1
 SQ SEQUENCE 15 AA; 1514 MW; 95E5C46F069C9775 CRC64;

Query Match 26.1%; Score 30; DB 10; Length 15;
 Best Local Similarity 50.0%; Pred. No. 2e+03;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 10 GAEVKRSAAEAK 21
 |||:|:|:|:|:
 Db 4 GAKVTKAAAKK 15

RESULT 10
 Q8K407
 ID Q8K407 PRELIMINARY; PRT; 17 AA.
 AC Q8K407;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Tissue-specific calpain 2 (fragment).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=22145722; PubMed=12150941;
RA Duan W.R., Ito M., Lee E.J., Chien P.-Y., Jameson J.L.;
RT "Estrogen regulates a tissue-specific calpain in the anterior
RT pituitary.";
RL Biochem. Biophys. Res. Commun. 295:261-266(2002).
DR EMBL; AF514419; AAM94284.1; -.
FT NON_TER 17
SQ SEQUENCE 17 AA; 1630 MW; 2027F409989F3A86 CRC64;

Query Match 26.1%; Score 30; DB 11; Length 17;
Best Local Similarity 53.3%; Pred. No. 2.2e+03;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 AAGAGAEVVKRSAAE 19
Db 2 AALAAAGVSKQRAVAE 16

RESULT 11
Q83877 PRELIMINARY; PRT; 24 AA.
AC Q83877;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE Putative ORF3 (Fragment).
OS Norwalk virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=11983;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SRSV-KY-89/89/J;
RX MEDLINE=94335115; PubMed=8057474;
RA Wang J., Jiang X., Madore H.P., Gray J., Desselberger U., Ando T.,
RA Seto Y., Oishi I., Lew J.F., Green K.Y., et al.;
RT "Sequence diversity of small, round-structured viruses in the Norwalk
RT virus group.";
RL J. Virol. 68:5982-5990(1994).
DR EMBL; L23828; AAA59230.1; -.
DR InterPro: IPR004278; RNA_capsid.
DR Pfam: PF03035; RNA_capsid; 1.
FT NON_TER 24
SQ SEQUENCE 24 AA; 2142 MW; 3F47CEE864CA957B CRC64;

Query Match 26.1%; Score 30; DB 12; Length 24;
Best Local Similarity 58.3%; Pred. No. 3.2e+03;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 TAPAAEGAGAEV 13
Db 13 TAGSALGAGIQV 24

RESULT 12
O55167 PRELIMINARY; PRT; 18 AA.
AC O55167;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ARE1 (Fragment).
GN ARE1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEW.1W/Gun;
RX MEDLINE=99009314; PubMed=9790748;

```

```

RA Walter L., Guenther E.;
RT "Identification of a novel highly conserved gene in the centromeric
RT part of the major histocompatibility complex.";
RL Genomics 52:298-304(1998).
DR EMBL; AJ223831; CAAL1568.1; -.
FT NON_TER 18
SQ SEQUENCE 18 AA; 1505 MW; B88156A2048C3388 CRC64;

Query Match 25.2%; Score 29; DB 11; Length 18;
Best Local Similarity 75.0%; Pred. No. 3.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 APAAEGAG 10
Db 8 AAAAQQAG 15

RESULT 13
Q9T217 PRELIMINARY; PRT; 22 AA.
AC Q9T217;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 14.3 photosystem I PSAE protein (Fragment).
OS Nicotiana sylvestris (Wood tobacco).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4096;
RN [1]
RP SEQUENCE.
RX MEDLINE=94105345; PubMed=8278548;
RA Obokata J., Mikami K., Hayashida N., Nakamura M., Sugiura M.;
RT "Molecular heterogeneity of photosystem I. psad, psae, psaf, psah, and
RT psal are all present in isoforms in Nicotiana spp.";
RL Plant Physiol. 102:1259-1267(1993).
SQ SEQUENCE 22 AA; 2033 MW; 488E17845223FFF7 CRC64;

Query Match 25.2%; Score 29; DB 8; Length 22;
Best Local Similarity 44.4%; Pred. No. 4e+03;
Matches 8; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 3 APAAEGAGAEVVKRSAAEA 20
Db 5 APPAAATAEPAEAPVKA 22

RESULT 14
O48861 PRELIMINARY; PRT; 22 AA.
AC O48861;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Homeobox protein (Fragment).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IRBB21;
RA Liu G., Yang J., Zhai W., He P., Li X., Lu J., Li S., Zhu L.;
RT "Mapping study of rice homeobox gene family";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003600; AAB94485.1; -.
DR Gramene; O48861; -.
FT NON_TER 1
FT NON_TER 22
SQ SEQUENCE 22 AA; 2337 MW; B18E95E5952F6CE9C CRC64;

```

Query Match 25.2%; Score 29; DB 10; Length 22;
Best Local Similarity 40.0%; Pred. No. 4e+03;
Matches 6; Conservative 2; Mismatches 0; Indels 7; Gaps 0;

QY 1 GTAPAAEGAGAEVKR 15
| | | | | : |
DB 1 GKAALAEETGLDAKQ 15

RESULT 15

QPT2I0
ID Q9T2I0 PRELIMINARY; PRT; 15 AA.
AC Q9T2I0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 18.5 kDa photosystem I PSAD protein (Fragment).
OS Nicotiana tomentosiformis (Tobacco).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4098;
RN [1]
RP SEQUENCE.
RX MEDLINE=94105345; PubMed=8278548;
RA Obokata J., Mikami K., Hayashida N., Nakamura M., Sugiura M.;
RT "Molecular heterogeneity of photosystem I: psad, psaE, psaF, psaH, and
RT psal are all present in isoforms in Nicotiana spp.";
RL Plant Physiol. 102:1259-1267 (1993).
DR PIR; PQ0692; PQ0692.
SQ SEQUENCE 15 AA; 1429 MW; D2388E488B5760A8 CRC64;

Query Match 24.3%; Score 28; DB 8; Length 15;
Best Local Similarity 50.0%; Pred. No. 3.7e+03;
Matches 7; Conservative 0; Mismatches 7; Indels 7; Gaps 0;

QY 5 AAEGAGAEVKRASA 18
| | | | | | |
DB 1 AEEAAAATKEAEA 14

Search completed: October 7, 2004, 17:56:37
Job time : 76 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:36:12 ; Search time 84.5 Seconds
(without alignments)
80.250 Million cell updates/sec

Title: US-10-802-644-2

Perfect score: 115

Sequence: 1 GTAPAAEGAGAEVKSASAKQAF 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 538975

Minimum DB seq length: 0

Maximum DB seq length: 24

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_29Jan04.*

1: geneseqp1980s.*

2: geneseqp1990s.*

3: geneseqp2000s.*

4: geneseqp2001s.*

5: geneseqp2002s.*

6: geneseqp2003as.*

7: geneseqp2003bs.*

8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	100.0	24	6	Abp97766 Scrambled
2	44	38.3	18	7	Add23833 Breast ca
3	38	33.0	24	5	Abp62492 Human imm
4	36	31.3	8	4	Abp12991 HIV A02 s
5	36	31.3	8	4	Abp20973 HIV A03 m
6	36	31.3	8	7	Add56874 HLA bindi
7	36	31.3	9	4	Abp16782 HIV B07 s
8	36	31.3	9	4	Abp17983 HIV B58 s
9	36	31.3	9	4	Abp13017 HIV A02 s
10	36	31.3	10	7	Add57289 HLA bindi
11	36	31.3	11	4	Abp13067 HIV A02 s
12	36	31.3	11	7	Add57290 HLA bindi
13	36	31.3	20	2	AAR43045 N-termina
14	36	31.3	22	2	AAR95258 Vector en
15	36	31.3	22	5	Aau99665 pET15b en
16	36	31.3	22	6	Abu09796 Fusion pr
17	36	31.3	22	6	Abg72306 Amino-ter
18	36	31.3	23	2	AAR41278 pET-3A de
19	35	30.4	15	4	AAG62627 Human RNA
20	35	30.4	24	5	Abp62462 Human imm
21	35	30.4	24	5	Abp62465 Human imm
22	35	30.4	24	5	Abp62463 Human imm
23	35	30.4	24	5	Abp62481 Human imm
24	35	30.4	24	5	Abp62464 Human imm
25	35	30.4	24	5	Abp62491 Human imm

26	35	30.4	24	5	ABP62480 Human imm
27	35	30.4	24	5	ABP62479 Human imm
28	35	30.4	24	5	ABP62488 Human imm
29	35	30.4	24	5	ABP62490 Human imm
30	34	29.6	14	6	ABP74859 Proteome
31	34	29.6	20	4	ABB76936 Rat Vg1-1
32	34	29.6	21	2	AAY20200 Human bet
33	34	29.6	23	6	AAE38088 Human COU
34	34	29.6	24	5	ABP62478 Human imm
35	33.5	29.1	10	4	AAG95102 Human com
36	33.5	29.1	10	4	AAG97627 Human com
37	33.5	29.1	24	5	ABG65791 Plant pot
38	33	28.7	8	4	ABP12990 HIV A02 s
39	33	28.7	8	4	ABP20942 HIV A03 m
40	33	28.7	9	4	ABP16781 HIV B07 s
41	33	28.7	9	4	ABP17977 HIV B58 s
42	33	28.7	9	4	ABP13016 HIV A02 s
43	33	28.7	10	4	ABP18930 HIV B62 s
44	33	28.7	10	4	ABP16789 HIV B07 s
45	33	28.7	10	7	ADD57307 HLA bindi

ALIGNMENTS

RESULT 1
ABP97766
ID ABP97766 standard; peptide; 24 AA.
XX
AC ABP97766;
XX
DT 11-AUG-2003 (first entry)
XX
DE Scrambled myristoylated N-terminal sequence (MANS) peptide.
XX
KW Myristoylated N-terminal sequence; MANS; MARCKS; mucus secretion;
KW inflammatory mediator; inflammation; respiratory disease; asthma;
KW chronic bronchitis; chronic obstructive pulmonary disease; COPD;
KW bowel disease; irritable bowel syndrome; Crohn's disease;
KW ulcerative colitis; skin disease; rosacea; psoriasis; acne;
KW autoimmune disease; pain; arthritis; cystic fibrosis.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal myristate chain attached"
XX
WO2003000027-A2.
XX
03-JAN-2003.
XX
26-JUN-2002; 2002WO-US022270.
XX
26-JUN-2001; 2001US-0300933P.
XX
(UYNC-) UNIV NORTH CAROLINA STATE.
XX
Martin LD, Adler KB, Li Y;
XX
WPI; 2003-278239/27.
XX
Method of regulating inflammation comprises administering a composition
comprising a MANS peptide or an active fragment thereof.
XX
Disclosure; Page 26; 54pp; English.
XX
The present sequence represents a scrambled myristoylated N-terminal
sequence (MANS) peptide. It is used as a control peptide, in the method
of the invention. The MANS peptide is identical to the first 24 amino
acids of MARCKS, and mediates insertion of MARCKS into membranes. The
MANS peptide inhibits both mucus secretion and inflammatory mediators.
XX
The specification describes a method of regulating inflammation. The

CC method comprises administering a composition comprising a MANS peptide.
CC The method is useful for treating inflammation caused by respiratory
CC diseases (e.g. asthma, chronic bronchitis and chronic obstructive
CC pulmonary disease (COPD), bowel diseases (e.g. irritable bowel syndrome,
CC Crohn's disease and ulcerative colitis), skin diseases (e.g. rosacea,
CC eczema, psoriasis and severe acne), autoimmune diseases and pain
CC syndromes, arthritis and cystic fibrosis

XX Sequence 24 AA;
SQ

Query Match 100.0%; Score 115; DB 6; Length 24;
Best local Similarity 100.0%; Pred. NO. 6.8e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTAPAAEGAGAEVVKRASAIAKQAF 24
| | | | | | | | | | | | | | | |
Db 1 GTAPAAEGAGAEVVKRASAIAKQAF 24
| | | | | | | | | | | | | | | |

RESULT 2
ID ADD23833 standard; peptide; 18 AA.
XX AC ADD23833;
XX DT 15-JAN-2004 (first entry)
XX DE Breast cancer membrane protein (BCMP) peptide SEQ ID NO:660.
XX KW breast cancer; screening; diagnosis; breast cancer therapy;
XX KW breast cancer membrane protein; BCMP; cytostatic; vaccine; human.
XX OS Homo sapiens.
XX FN WO2003087831-A2.
XX PD 23-OCT-2003.
XX PP 10-APR-2003; 2003WO-GB001559.
XX PR 11-APR-2002; 2002GB-00008331.
XX PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX PI Hudson LJ, Stamps AC, Terrett JA;
XX DR WPI; 2003-845381/78.

Screening, diagnosing and/or treating breast cancer by detecting a change
in expression or activity of a breast cancer membrane protein (BCMP)
polypeptide or encoding nucleic acid molecule.

PS Example; SEQ ID NO 660; 81pp; English.

The present invention describes a method of screening for and/or
diagnosing breast cancer in a subject, and/or monitoring the
effectiveness of breast cancer therapy. The method comprises detecting
and/or quantifying in a biological sample obtained from the subject a
breast cancer membrane protein (BCMP) polypeptide and a nucleic acid
molecule. Also described: (1) an antibody, its functionally-active
fragment, derivative or analogue, that specifically binds to one or more
of the BCMP polypeptide; (2) a diagnostic kit comprising a capture
reagent specific for an BCMP polypeptide, reagents and instructions for
use; (3) a method for screening for anti-breast cancer agents that
interact with the BCMP polypeptide, comprising contacting the polypeptide
with a candidate agent, and determining whether or not the candidate
agent interacts with the polypeptide; (4) a method for screening for anti-
breast cancer agents that modulate the expression or activity of an BCMP
polypeptide or the nucleic acid molecule cited above, comprising
comparing the expression or activity of the polypeptide or nucleic acid
molecule, in the presence and absence of a candidate agent or in the
presence of a control agent, and determining whether the candidate agent
causes the expression or activity of the polypeptide or nucleic acid

CC	molecule to change; and (5) an agent identified by the method of (3) or
CC	(4), which interacts with the polypeptide or causes the expression or
CC	activity of the polypeptide, or the expression of the nucleic acid
CC	molecule to change. BCPMs have cytotaxtic activities, and can be used in
CC	vaccines. The BCPM polypeptide, nucleic acid molecule, antibody, agent or
CC	their derivatives, are useful in the manufacture of a medicament for the
CC	treatment of breast cancer, where the composition is a vaccine. The
CC	present sequence represents a BCPM peptide which is used in the
CC	exemplification of the present invention.
XX	
SQ	Sequence 18 AA;
Query Match	38.3%; Score 44; DB 7; Length 18;
Best Local Similarity	53.3%; Pred. No. 14;
Matches	8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
Qy	1 GTAPAAEGAGAEVKR 15 : :
Db	4 GSAPAGEGEGVTLQR 18 : :
RESULT 3	
ABP62492	ID ABP62492 standard; peptide; 24 AA.
XX	AC ABP62492;
XX	DT 10-OCT-2002 (first entry)
XX	Human immunopeptide to HCV E2 glycoprotein framework sequence #31.
XX	Virucide; human; immunopolypeptide; immunopeptide; envelope glycoprotein;
KW	nonstructural protein; hepatitis C virus; HCV; E2 glycoprotein;
KW	NS3 protein; viral infection.
OS	Homo sapiens.
XX	WO200259340-A1.
PN	01-AUG-2002.
PD	25-JAN-2002; 2002WO-US002303.
XX	26-JAN-2001; 2001US-0264451P.
PR	(SCRI) SCRIPPS RES INST.
PA	Mariyama T, Jones IM, Burton DR, Fox RI;
PI	WPI; 2002-599801/64.
XX	New human immunopolypeptide with binding specificity for certain envelope
PT	glycoproteins and nonstructural proteins of hepatitis C virus (HCV), for
PT	diagnosing or treating patients having or suspected of having HCV
PT	infection.
XX	Claim 4; Fig 17; 308pp; English.
PS	The present invention relates to human immunopolypeptides, produced by a
CC	phage transfected cell library. The present sequence is one such
CC	immunopolypeptide. The immunopolypeptides have binding specificity for
CC	envelope glycoprotein E2 and nonstructural protein NS3 of hepatitis C
CC	virus (HCV). E2 glycoprotein is believed to be responsible for target
CC	cell binding and contains neutralising epitopes, while NS3 is thought to
CC	be involved in the replication of HCV. The immunopolypeptides are useful
CC	for diagnosing and treating a patient having or suspected to be having
CC	HCV infection
XX	
SQ	Sequence 24 AA;
Query Match	33.0%; Score 38; DB 5; Length 24;
Best Local Similarity	53.3%; Pred. No. 1.5e+02;
Matches	8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 7 EGAGAEVKKRASAEAK 21
 Db :|||||:
 2 EQSGAEVKKPGASVK 16

RESULT 4
 ABP12991
 ID ABP12991 standard; peptide; 8 AA.

XX AC ABP12991;
 XX DT 11-SEP-2003 (revised)
 DT 15-JUL-2002 (first entry)
 XX DE HIV A02 super motif nef peptide #4.

XX KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
 KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
 KW vaccine; HIV infection; immunisation; virucide.

XX OS Human immunodeficiency virus 1.

XX PN WO200124810-A1.

XX PD 12-APR-2001.

XX PF 05-OCT-2000; 2000WO-US027766.

XX PR 05-OCT-1999; 99US-00412863.

XX PA (EPIM-) EPIMMUNE INC.

XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
 PI Baker DM, Celis E, Kubo RT, Grey HM;

XX WPI; 2001-354887/37.

XX Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
 PT peptide groups, useful for vaccinating against HIV-1.

XX Claim 32; Page 134; 448pp; English.

XX The present invention describes a composition (I) comprising a prepared
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
 CC be used for immunising subjects against HIV-1 infections. The use of
 CC group-based vaccines has several advantages over traditional vaccines,
 CC particularly when compared to the use of whole antigens in vaccine
 CC compositions. There is evidence that the immune response to whole
 CC antigens is directed largely toward variable regions of the antigen,
 CC allowing for immune escape due to mutations. The groups for inclusion in
 CC an group-based vaccine may be selected from conserved regions of viral or
 CC tumour-associated antigens, which therefore reduces the likelihood of
 CC escape mutants. Furthermore, immunosuppressive groups that may be present
 CC in whole antigens can be avoided with the use of group-based vaccines. An
 CC additional advantage of an group-based vaccine approach is the ability to
 CC combine selected groups (CTL and HTL), and further, to modify the
 CC composition of the groups, achieving, for example, enhanced
 CC immunogenicity. Accordingly, the immune response can be modulated, as
 CC appropriate, for the target disease. Similar engineering of the response
 CC is not possible with traditional approaches. ABP11501 to ABP25412
 CC represent peptide sequences used in the exemplification of the present
 CC invention. (Updated on 11-SEP-2003 to standardise OS field)

XX Sequence 8 AA;

Query Match 31.3%; Score 36; DB 4; Length 8;
 Best Local Similarity 87.5%; Pred. No. 1.4e+06;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAAGAGAGA 11

Db :|||||:
 1 PAAEGVGA 8

RESULT 5
 ABP20973
 ID ABP20973 standard; peptide; 8 AA.

XX AC ABP20973;

XX DT 11-SEP-2003 (revised)
 DT 15-JUL-2002 (first entry)

XX DE HIV A03 motif nef peptide #123.

XX KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
 KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
 KW vaccine; HIV infection; immunisation; virucide.

XX OS Human immunodeficiency virus 1.

XX PN WO200124810-A1.

XX PD 12-APR-2001.

XX PF 05-OCT-2000; 2000WO-US027766.

XX PR 05-OCT-1999; 99US-00412863.

XX PA (EPIM-) EPIMMUNE INC.

XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
 PI Baker DM, Celis E, Kubo RT, Grey HM;

XX WPI; 2001-354887/37.

XX Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
 PT peptide groups, useful for vaccinating against HIV-1.

XX Claim 32; Page 299; 448pp; English.

XX The present invention describes a composition (I) comprising a prepared
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
 CC be used for immunising subjects against HIV-1 infections. The use of
 CC group-based vaccines has several advantages over traditional vaccines,
 CC particularly when compared to the use of whole antigens in vaccine
 CC compositions. There is evidence that the immune response to whole
 CC antigens is directed largely toward variable regions of the antigen,
 CC allowing for immune escape due to mutations. The groups for inclusion in
 CC an group-based vaccine may be selected from conserved regions of viral or
 CC tumour-associated antigens, which therefore reduces the likelihood of
 CC escape mutants. Furthermore, immunosuppressive groups that may be present
 CC in whole antigens can be avoided with the use of group-based vaccines. An
 CC additional advantage of an group-based vaccine approach is the ability to
 CC combine selected groups (CTL and HTL), and further, to modify the
 CC composition of the groups, achieving, for example, enhanced
 CC immunogenicity. Accordingly, the immune response can be modulated, as
 CC appropriate, for the target disease. Similar engineering of the response
 CC is not possible with traditional approaches. ABP11501 to ABP25412
 CC represent peptide sequences used in the exemplification of the present
 CC invention. (Updated on 11-SEP-2003 to standardise OS field)

XX Sequence 8 AA;

Query Match 31.3%; Score 36; DB 4; Length 8;
 Best Local Similarity 87.5%; Pred. No. 1.4e+06;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAAEGVGA 11

Db :|||||:
 1 PAAEGVGA 8

DT 15-JUL-2002 (first entry)
 XX HIV B58 super motif nef peptide #19.
 DE HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
 KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
 KW vaccine; HIV infection; immunisation; virucide.
 XX Human immunodeficiency virus 1.
 OS WO200124810-A1.
 XX PN 12-APR-2001.
 XX PD 05-OCT-2000; 2000WO-US027766.
 XX PF 05-OCT-1999; 99US-00412863.
 XX PR (EPTM-) EPIMUNE INC.
 XX PA Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
 XX PI Baker DM, Celis E, Kubo RT, Grey HM;
 XX PT WPI; 2001-354887/37.
 XX DR Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
 XX PT peptide groups, useful for vaccinating against HIV-1.
 XX PS Claim 32; Page 237; 448pp; English.
 XX CC The present invention describes a composition (I) comprising a prepared
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
 CC be used for immunising subjects against HIV-1 infections. The use of
 CC group-based vaccines has several advantages over traditional vaccines,
 CC particularly when compared to the use of whole antigens in vaccine
 CC compositions. There is evidence that the immune response to whole
 CC antigens is directed largely toward variable regions of the antigen,
 CC allowing for immune escape due to mutations. The groups for inclusion in
 CC an group-based vaccine may be selected from conserved regions of viral or
 CC tumour-associated antigens, which therefore reduces the likelihood of
 CC escape mutants. Furthermore, immunosuppressive groups that may be present
 CC in whole antigens can be avoided with the use of group-based vaccines. An
 CC additional advantage of an group-based vaccine approach is the ability to
 CC combine selected groups (CTL and HTL), and further, to modify the
 CC composition of the groups, achieving, for example, enhanced
 CC immunogenicity. Accordingly, the immune response can be modulated, as
 CC appropriate, for the target disease. Similar engineering of the response
 CC is not possible with traditional approaches. ABP11501 to ABP25412
 CC represent peptide sequences used in the exemplification of the present
 CC invention. (Updated on 11-SEP-2003 to standardise OS field)
 XX SQ Sequence 9 AA;

Query Match 31.3%; Score 36; DB 4; Length 9;
 Best Local Similarity 87.5%; Pred. No. 1.4e+06;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAEEGAGA 11
 |||||
 Db 1 PAAEGVGA 8

RESULT 9
 ABP13017
 ID ABP13017 standard; peptide; 9 AA.
 XX AC ABP13017;
 XX 11-SEP-2003 (revised)
 DT 15-JUL-2002 (first entry)
 XX

DE HIV A02 super motif nef peptide #30.
 XX HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
 KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
 KW vaccine; HIV infection; immunisation; virucide.
 XX Human immunodeficiency virus 1.
 OS WO200124810-A1.
 XX PN 12-APR-2001.
 XX PD 05-OCT-2000; 2000WO-US027766.
 XX PF 05-OCT-1999; 99US-00412863.
 XX PR (EPTM-) EPIMUNE INC.
 XX PA Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
 XX PI Baker DM, Celis E, Kubo RT, Grey HM;
 XX PT WPI; 2001-354887/37.
 XX DR Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
 XX PT peptide groups, useful for vaccinating against HIV-1.
 XX PS Claim 32; Page 135; 448pp; English.
 XX CC The present invention describes a composition (I) comprising a prepared
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
 CC be used for immunising subjects against HIV-1 infections. The use of
 CC group-based vaccines has several advantages over traditional vaccines,
 CC particularly when compared to the use of whole antigens in vaccine
 CC compositions. There is evidence that the immune response to whole
 CC antigens is directed largely toward variable regions of the antigen,
 CC allowing for immune escape due to mutations. The groups for inclusion in
 CC an group-based vaccine may be selected from conserved regions of viral or
 CC tumour-associated antigens, which therefore reduces the likelihood of
 CC escape mutants. Furthermore, immunosuppressive groups that may be present
 CC in whole antigens can be avoided with the use of group-based vaccines. An
 CC additional advantage of an group-based vaccine approach is the ability to
 CC combine selected groups (CTL and HTL), and further, to modify the
 CC composition of the groups, achieving, for example, enhanced
 CC immunogenicity. Accordingly, the immune response can be modulated, as
 CC appropriate, for the target disease. Similar engineering of the response
 CC is not possible with traditional approaches. ABP11501 to ABP25412
 CC represent peptide sequences used in the exemplification of the present
 CC invention. (Updated on 11-SEP-2003 to standardise OS field)
 XX SQ Sequence 9 AA;

Query Match 31.3%; Score 36; DB 4; Length 9;
 Best Local Similarity 87.5%; Pred. No. 1.4e+06;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAAEGAGA 11
 |||||
 Db 1 PAAEGVGA 8

RESULT 10
 ADD57289
 ID ADD57289 standard; peptide; 10 AA.
 XX AC ADD57289;
 XX 15-JAN-2004 (first entry)
 DT HLA binding epitope #509.
 DE Virucide; cytostatic; gene therapy; vaccine; epitope; cytotoxic T cell;
 KW

KW MHC class I; CTL; HTL; A2-restricted cytotoxic lymphocyte; HLA;
KW viral disease; cancer.
XX Unidentified.
XX WO2003040165-A2.
XX 15-MAY-2003.
XX 18-OCT-2001; 2001WO-US051650.
XX 19-OCT-2000; 2000US-0242350P.
XX 20-APR-2001; 2001US-0285624P.
XX (EPIM-) EPIMUNE INC.
XX Sette A, Sidney J, Southwood S;
XX WPI; 2003-441519/41.
XX New composition comprising at least one peptide having allele-specific
XX binding motifs for HLA, useful for preventing, treating or diagnosing
XX viral diseases and cancer.
XX Claim 1; Page 52-379; 382pp; English.
XX The invention relates to a composition comprising at least one peptide
XX having an isolated, prepared epitope selected from any of the sequences
XX from 30 lists given in the specification. Also disclosed is a method for
XX inducing a cytotoxic T cell response against a pre-selected antigen in a
XX patient expressing a specific MHC class I allele by contacting cytotoxic
XX T cells from the patient with the composition cited above. The
XX composition comprises an epitope that is joined by an amino acid linker.
XX The epitope is admixed or joined to a CTL or HTL epitope. The epitope is
XX bound to an HLA molecule on the antigen-presenting cell, where when an A2
XX -restricted cytotoxic lymphocyte (CTL) is present, a receptor of the CTL
XX binds to a complex of the HLA molecule and the epitope. Specifically
XX claimed are peptides having allele-specific binding motifs for HLA. The
XX compositions and methods are useful for preventing, treating or
XX diagnosing viral diseases and cancer. The peptide epitopes are useful as
XX diagnostic agents for evaluating immune responses, for making antibodies
XX and for evaluating efficacy of a vaccine. Sequences given in ADD56781-
XX ADD65275 represent epitopes of the invention as given in Tables 2-31.
XX Sequence 10 AA;
SQ
Query Match 31.3%; Score 36; DB 7; Length 10;
Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 PAAEGAGA 11
DB 3 PAAEGVGA 10
RESULT 11
ABP13067
ID ABP13067 standard; peptide; 11 AA.
XX ABP13067;
XX 11-SEP-2003 (revised)
XX 15-JUL-2002 (first entry)
XX HIV A02 super motif nef peptide #80.
XX HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
KW vaccine; HIV infection; immunisation; virucide.
XX Human immunodeficiency virus 1.
XX WO200124810-A1.

XX 12-APR-2001.
XX 05-OCT-2000; 2000WO-US027766.
XX 05-OCT-1999; 99US-00412863.
XX (EPIM-) EPIMUNE INC.
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX Baker DM, Cellis E, Kubo RT, Grey HM;
XX WPI; 2001-354887/37.
XX Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
XX peptide groups, useful for vaccinating against HIV-1.
XX Claim 32; Page 136; 448pp; English.
XX The present invention describes a composition (I) comprising a prepared
XX human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
XX sequence selected from 51 defined amino acid sequences (ABL25347 to
XX ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
XX be used for immunising subjects against HIV-1 infections. The use of
XX group-based vaccines has several advantages over traditional vaccines,
XX particularly when compared to the use of whole antigens in vaccine
XX compositions. There is evidence that the immune response to whole
XX antigens is directed largely toward variable regions of the antigen,
XX allowing for immune escape due to mutations. The groups for inclusion in
XX an group-based vaccine may be selected from conserved regions of viral or
XX tumour-associated antigens, which therefore reduces the likelihood of
XX escape mutants. Furthermore, immunosuppressive groups that may be present
XX in whole antigens can be avoided with the use of group-based vaccines. An
XX additional advantage of an group-based vaccine approach is the ability to
XX combine selected groups (CTL and HTL), and further, to modify the
XX composition of the groups, achieving, for example, enhanced
XX immunogenicity. Accordingly, the immune response can be modulated, as
XX appropriate, for the target disease. Similar engineering of the response
XX is not possible with traditional approaches. ABP11501 to ABP25412
XX represent peptide sequences used in the exemplification of the present
XX invention. (Updated on 11-SEP-2003 to standardise OS field)
XX Sequence 11 AA;
SQ
Query Match 31.3%; Score 36; DB 4; Length 11;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 PAAEGAGA 11
DB 4 PAAEGVGA 11
RESULT 12
ADD57290
ID ADD57290 standard; peptide; 11 AA.
XX ADD57290;
XX 15-JAN-2004 (first entry)
XX HLA binding epitope #510.
XX Virucide; cytostatic; gene therapy; vaccine; epitope; cytotoxic T cell;
KW MHC class I; CTL; HTL; A2-restricted cytotoxic lymphocyte; HLA;
KW viral disease; cancer.
XX Unidentified.
XX WO2003040165-A2.
XX 15-MAY-2003.
XX

PF 18-OCT-2001; 2001WO-US051650.
 XX
 PR 19-OCT-2000; 2000US-0242350P.
 PR 20-APR-2001; 2001US-0285624P.
 XX
 PA (EPIM-) EPIMUNE INC.
 XX
 PI Sette A, Sidney J, Southwood S;
 XX WPI; 2003-441519/41.
 XX
 XX New composition comprising at least one peptide having allele-specific
 PT binding motifs for HLA, useful for preventing, treating or diagnosing
 PT viral diseases and cancer.
 XX
 XX Claim 1; Page 52-379; 382pp; English.
 PS
 PS The invention relates to a composition comprising at least one peptide
 CC having an isolated, prepared epitope selected from any of the sequences
 CC from 30 lists given in the specification. Also disclosed is a method for
 CC inducing a cytotoxic T cell response against a pre-selected antigen in a
 CC patient expressing a specific MHC class I allele by contacting cytotoxic
 CC T cells from the patient with the composition cited above. The
 CC composition comprises an epitope that is joined by an amino acid linker.
 CC The epitope is admixed or joined to a CTL or HTL epitope. The epitope is
 CC bound to an HLA molecule on the antigen-presenting cell, where when an A2
 CC -restricted cytotoxic lymphocyte (CTL) is present, a receptor of the CTL
 CC binds to a complex of the HLA molecule and the epitope. Specifically
 CC claimed are peptides having allele-specific binding motifs for HLA. The
 CC compositions and methods are useful for preventing, treating or
 CC diagnosing viral diseases and cancer. The peptide epitopes are useful as
 CC diagnostic agents for evaluating immune responses, for making antibodies
 CC and for evaluating efficacy of a vaccine. Sequences given in ADP56781-
 CC ADD65275 represent epitopes of the invention as given in Tables 2-31.
 XX
 SQ Sequence 11 AA;
 Query Match 31.3%; Score 36; DB 7; Length 11;
 Best Local Similarity 87.5%; Pred. No. 1.3e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 PAAEGAGA 11
 Db 3 PAAEGVGA 10
 RESULT 13
 AAR43045
 ID AAR43045 standard; protein; 20 AA.
 XX
 AC AAR43045;
 XX
 DT 25-MAR-2003 (revised)
 DT 10-MAR-2003 (revised)
 DT 24-MAY-1994 (first entry)
 XX
 XX N-terminal fragment.
 DE
 DE Asparaginylendopeptidase; Canavalia ensiformis; seed; L-asparagine;
 KW primer; PCR; protein fragmentation; peptide synthesis.
 KW
 OS Canavalia ensiformis.
 XX
 XX JP05276960-A.
 XX
 XX 26-OCT-1993.
 XX
 XX 07-AUG-1992; 92JP-00231602.
 XX
 XX 07-FEB-1992; 92JP-00056023.
 XX
 XX (TAKI) TAKARA SHUZO CO LTD.
 PA
 XX

DR WPI; 1993-373587/47.
 XX
 PT New gene for encoding asparaginyl endo-peptidase - comprises 8 specified
 PT DNA sequences.
 XX
 XX Disclosure; Page 30; 35pp; Japanese.
 XX
 CC A gene encoding asparaginylendopeptidase is claimed. 8 DNA sequences are
 CC given (AAQ50559-66). The enzyme is a protease derived from a seed of
 CC Canavalia ensiformis which selectively hydrolyses C-terminus amide bond
 CC of L-asparagine residue (see AAR43033 and AAR43041). The enzyme is useful
 CC for protein fragmentation and enzymatic peptide synthesis. The primers
 CC given in AAQ50567-68, AAQ50576-77 and AAQ50583-90 were used in the
 CC isolation of the fragments given in AAQ50569-75 and AAQ50578-79, by PCR.
 CC (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003
 CC to correct PA field.)
 XX
 SQ Sequence 20 AA;
 Query Match 31.3%; Score 36; DB 2; Length 20;
 Best Local Similarity 47.4%; Pred. No. 2.4e+02;
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
 QY 4 PAAEGAGAEVKRAAEAKQ 22
 Db 2 PAAKARKEAEELAAATAEQ 20
 RESULT 14
 AAR95258
 ID AAR95258 standard; protein; 22 AA.
 XX
 AC AAR95258;
 XX
 DT 09-NOV-1996 (first entry)
 XX
 DE Vector encoded amino acids used to increase protease yield.
 XX
 KW Protease; restriction protease; cleavage; proteolytic domain; protein.
 XX
 OS Synthetic.
 XX
 XX WO9616167-A1.
 XX
 PD 30-MAY-1996.
 XX
 PF 16-NOV-1995; 95WO-N0000213.
 XX
 PR 17-NOV-1994; 94NO-00004411.
 XX
 PA (HAVA/) HAVARSTEIN L S.
 PA (NESI/) NES I F.
 XX
 PI Havarstein LS, Nes IF;
 XX
 DR WPI; 1996-268604/27.
 XX
 PT Protease(s) with highly specific cleavage activity - have proteolytic
 PT domains derived from N-terminus of ABC transporters, pref. 150 N-terminal
 PT amino acids of LagD.
 XX
 XX Claim 10; Page 22; 43pp; English.
 PS
 CC Proteases derived from ABC-transporters containing N-terminal proteolytic
 CC domains are restriction proteases that can be used for the cleavage of
 CC proteins at specific amino acid cleavage sites. These 22 vector encoded
 CC amino acids were added to the C-terminus of the subcloned protease
 CC sequence to substantially increase the yield of the recombinant ABC
 CC protease whilst not reducing its enzymatic activity of the protease
 XX
 SQ Sequence 22 AA;
 Query Match 31.3%; Score 36; DB 2; Length 22;

Best Local Similarity 47.4%; Pred. No. 2.7e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 PAAEAGAGAEVKRASAEAKQ 22
||| | : | : | : |
Db 4 PAANKARKEAELEAAATAEQ 22

CC interfere specifically with the interaction between the D1 polypeptide of
CC the human CAR protein and a viral attachment protein which binds to the
CC D1 polypeptide. D1 is useful for structural determination of virus
CC attachment proteins (e.g., adenovirus 12 (Ad12) knob, complexed to D1 or
CC D1/D2) which facilitates the identification of residues that form the
CC receptor-knob interface. The gene for human CAR is located on chromosome
CC 21. The present sequence is the pET15b encoded C-terminal extension
CC peptide
XX
SQ Sequence 22 AA;

Query Match 31.3%; Score 36; DB 5; Length 22;
Best Local Similarity 47.4%; Pred. No. 2.7e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 PAAEAGAGAEVKRASAEAKQ 22
||| | : | : | : |
Db 4 PAANKARKEAELEAAATAEQ 22

RESULT 15
AAU99665
ID AAU99665 standard; peptide; 22 AA.
XX
AC AAU99665;
XX
DT 07-OCT-2002 (first entry)
XX
DE pET15b encoded C-terminal extension peptide.
XX
KW pET15b; C terminal extension; adenovirus binding domain; D1; D2; sd1;
KW coxsackievirus and adenovirus receptor; CAR; protein therapy; virucide;
KW adenovirus infection; coxsackievirus infection; ocular infection;
KW upper respiratory tract infection; gastrointestinal infection;
KW chromosome 21; virus attachment protein; Adenovirus 12; knob;
KW receptor-knob interface.
XX
OS Synthetic.
XX
PN US6395875-B1.
XX
PD 28-MAY-2002.
XX
PF 25-JAN-1999; 99US-00236423.
XX
PR 25-JAN-1999; 99US-00236423.
XX
PA (BROO-) BROOKHAVEN SCI ASSOC LLC.
XX
PI Freimuth PI;
XX
DR WPI; 2002-556105/59.
XX

Search completed: October 7, 2004, 17:53:34
Job time : 88.5 secs

New polypeptide for preparing a composition to treat ocular, upper
respiratory or gastrointestinal infection, comprises an adenovirus
binding domain D1 of human coxsackievirus and an adenovirus receptor
protein.

Claim 5; Fig 1; 13pp; English.

The invention relates to an isolated polypeptide consisting essentially
of an amino acid sequence which corresponds to adenovirus binding domain
D1 of human coxsackievirus and adenovirus receptor (CAR) protein. Also
included are an isolated polypeptide consisting essentially of an amino
acid sequence which corresponds to extracellular domains D1 and D2 of
human CAR protein, where the amino acid sequence corresponds to residues
20-237 of human pre-CAR protein with the exception of two amino acid
substitutions, Leu20Met and Ser21Gly and an isolated fusion protein (sd1)
comprising a first amino acid sequence which corresponds to amino acids
20-144 of human pre-CAR protein (D1) fused in frame to a second amino
acid sequence comprising the peptide appearing as AAU99665 where the
second amino acid sequence facilitates folding of D1 into a functional,
soluble domain when expressed in bacteria and is located C-terminal to
the first amino acid sequence, being encoded by the vector pET15b. CAR D1
functions as an antiviral agent by inhibiting viral infection of a cell.
CAR D1 is useful for treating a patient infected with virus e.g.,
adenovirus of group A, adenovirus of group C and coxsackievirus subgroup
B. D1 is also useful for inhibiting the progression or spread of
infection and for the preparation of a therapeutic composition for
treating an ocular, upper respiratory or gastrointestinal infection,
respectively. D1 is also useful for identifying and characterising
molecules which bind CAR through the D1 domain to study the infection
process and to develop new therapeutics. D1 is also utilised in binding
assays for identification and further characterisation of D1. D1 is also
utilised in binding assays to screen for antiviral compounds which

Result No.	Score	Query #		Length	DB	ID	Description
		Match					
1	115	100.0	24	12	US-10-180-753-2		Sequence 2, Appli
2	115	100.0	24	16	US-10-602-644-2		Sequence 2, Appli
3	115	33.0	21	15	US-10-313-986-547		Sequence 547, App
4	36	31.3	22	13	US-10-155-282-1		Sequence 1, Appli
5	36	31.3	22	14	US-10-199-290-1		Sequence 1, Appli
6	36	31.3	22	14	US-10-218-419-1		Sequence 1, Appli
7	36	31.3	22	14	US-10-037-243-20		Sequence 20, Appl
8	36	31.3	22	14	US-10-037-243-42		Sequence 42, Appl
9	36	31.3	23	14	US-10-037-243-4		Sequence 4, Appli
10	34	29.6	14	15	US-10-394-980-142		Sequence 142, App
11	34	29.6	20	14	US-10-029-386-27605		Sequence 27605, A
12	33.5	29.1	10	10	US-09-572-404B-1296		Sequence 1296, Ap
13	33.5	29.1	10	10	US-09-572-404B-3822		Sequence 3822, Ap
14	33	28.7	9	8	US-08-344-924-276		Sequence 276, App
15	33	28.7	14	10	US-09-624-438-16		Sequence 16, Appl

SUMMARIES

; Sequence 2, Application US/10802644

```
Publication No. US20040180836A1
; GENERAL INFORMATION:
; APPLICANT: Martin, Linda
; APPLICANT: Adler, Kenneth
; APPLICANT: Li, Yuehua
; TITLE OF INVENTION: BLOCKING PEPTIDE FOR INFLAMMATORY CELL SECRETION
; FILE REFERENCE: 5051.574
; CURRENT APPLICATION NUMBER: US/10/802,644
; CURRENT FILING DATE: 2004-03-17
; PRIOR APPLICATION NUMBER: US/10/180,753
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: N-terminal myristoylation
US-10-802-644-2

Query Match      100.0%; Score 115; DB 16; Length 24;
Best Local Similarity 100.0%; Pred. No. 4.3e+09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTAPAAEGAGAEVVKRASAERAKQAF 24
Db 1 GTAPAAEGAGAEVVKRASAERAKQAF 24

RESULT 3
US-10-313-986-547
; Sequence 547, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 547
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-986-547

Query Match      33.0%; Score 38; DB 15; Length 21;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 9 AGAEVVKRASAERAKQA 23
Db 7 AGASIKIAPAEAPDA 21

RESULT 4
US-10-155-282-1
; Sequence 1, Application US/10155282
; Publication No. US20020151686A1
; GENERAL INFORMATION:
; APPLICANT: Freimuth, Paul
; TITLE OF INVENTION: RECOMBINANT SOLUBLE ADENOVIRUS RECEPTOR
```

```
FILE REFERENCE: BNL-2007
; CURRENT APPLICATION NUMBER: US/10/155,282
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US/09/236,423
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: CLONING VECTOR
; OTHER INFORMATION: ENCODED SEQUENCES
US-10-155-282-1

Query Match      31.3%; Score 36; DB 13; Length 22;
Best Local Similarity 47.4%; Pred. No. 3.7e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 4 PAAEGAGAEVVKRASAERAKQ 22
Db 4 PAANKARKEAEELAAATAEQ 22

RESULT 5
US-10-199-290-1
; Sequence 1, Application US/10199290
; Publication No. US20030027338A1
; GENERAL INFORMATION:
; APPLICANT: Brookhaven Science Associates
; APPLICANT: Freimuth, Paul I
; TITLE OF INVENTION: Structure of Adenovirus Bound to Cellular Receptor CAR
; FILE REFERENCE: BSA 02-24
; CURRENT APPLICATION NUMBER: US/10/199,290
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/389,603
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 09/236,423
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cloning vector encoded sequence
; OTHER INFORMATION: nces
US-10-199-290-1

Query Match      31.3%; Score 36; DB 14; Length 22;
Best Local Similarity 47.4%; Pred. No. 3.7e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 4 PAAEGAGAEVVKRASAERAKQ 22
Db 4 PAANKARKEAEELAAATAEQ 22

RESULT 6
US-10-218-419-1
; Sequence 1, Application US/10218419
; Publication No. US20030077813A1
; GENERAL INFORMATION:
; APPLICANT: Brookhaven Science Associates
; APPLICANT: Freimuth, Paul I
; TITLE OF INVENTION: Structure of Adenovirus Bound to Cellular Receptor CAR
; FILE REFERENCE: BSA 02-24
; CURRENT APPLICATION NUMBER: US/10/218,419
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 09/389,603
; PRIOR FILING DATE: 1999-09-03
```



```
/
; PRIOR APPLICATION NUMBER: 09/236,423
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cloning vector encoded sequence
; OTHER INFORMATION: nces
US-10-218-419-1

Query Match      31.3%; Score 36; DB 14; Length 22;
Best Local Similarity 47.4%; Pred. No. 3.7e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 PAAEGAGAEVVKRSAAEAKQ 22
   ||| | | | | | | | | | | | | | | |
Db 4 PAANKARKEAEALAAATAEQ 22

RESULT 7
US-10-037-243-20
; Sequence 20, Application US/10037243
; Publication No. US20030134352A1
; GENERAL INFORMATION:
; APPLICANT: Brookhaven Science Associates, LLC.
; APPLICANT: Freimuth, Paul I
; APPLICANT: Zhang, Yian-Biao
; APPLICANT: Howitt, Jason A
; TITLE OF INVENTION: Facilitating Protein Folding and Solubility by Use of Peptide Ext
; FILE REFERENCE: BSA 01-22
; CURRENT APPLICATION NUMBER: US/10/037,243
; CURRENT FILING DATE: 2002-06-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic T7A peptide
US-10-037-243-20

Query Match      31.3%; Score 36; DB 14; Length 22;
Best Local Similarity 47.4%; Pred. No. 3.7e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 PAAEGAGAEVVKRSAAEAKQ 22
   ||| | | | | | | | | | | | | | | |
Db 4 PAANKARKEAEALAAATAEQ 22

RESULT 8
US-10-037-243-42
; Sequence 42, Application US/10037243
; Publication No. US20030134352A1
; GENERAL INFORMATION:
; APPLICANT: Brookhaven Science Associates, LLC.
; APPLICANT: Freimuth, Paul I
; APPLICANT: Zhang, Yian-Biao
; APPLICANT: Howitt, Jason A
; TITLE OF INVENTION: Facilitating Protein Folding and Solubility by Use of Peptide Ext
; FILE REFERENCE: BSA 01-22
; CURRENT APPLICATION NUMBER: US/10/037,243
; CURRENT FILING DATE: 2002-06-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```
/
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Encoded amino acids
US-10-037-243-42

Query Match      31.3%; Score 36; DB 14; Length 22;
Best Local Similarity 47.4%; Pred. No. 3.7e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 PAAEGAGAEVVKRSAAEAKQ 22
   ||| | | | | | | | | | | | | | | |
Db 4 PAANKARKEAEALAAATAEQ 22

RESULT 9
US-10-037-243-4
; Sequence 4, Application US/10037243
; Publication No. US20030134352A1
; GENERAL INFORMATION:
; APPLICANT: Brookhaven Science Associates, LLC.
; APPLICANT: Freimuth, Paul I
; APPLICANT: Zhang, Yian-Biao
; APPLICANT: Howitt, Jason A
; TITLE OF INVENTION: Facilitating Protein Folding and Solubility by Use of Peptide Ext
; FILE REFERENCE: BSA 01-22
; CURRENT APPLICATION NUMBER: US/10/037,243
; CURRENT FILING DATE: 2002-06-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide
US-10-037-243-4

Query Match      31.3%; Score 36; DB 14; Length 23;
Best Local Similarity 47.4%; Pred. No. 3.9e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 PAAEGAGAEVVKRSAAEAKQ 22
   ||| | | | | | | | | | | | | | | |
Db 5 PAANKARKEAEALAAATAEQ 23

RESULT 10
US-10-394-980-142
; Sequence 142, Application US/10394980
; Publication No. US20040005633A1
; GENERAL INFORMATION:
; APPLICANT: Vandekerckhove, Joel
; APPLICANT: Gevaert, Kris
; TITLE OF INVENTION: METHODS AND APPARATUS FOR GEL-FREE QUALITATIVE AND
; FILE REFERENCE: VBV-001
; CURRENT APPLICATION NUMBER: US/10/394,980
; CURRENT FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: PCT/EP02/03368
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US60/278,171
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US60/318,749
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US60/323,999
; PRIOR FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 473
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 142
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
```

; OTHER INFORMATION: part of MK01_HUMAN (P28482, mitogen-activated protein kinase 1)
US-10-394-980-142

Query Match 29.6%; Score 34; DB 15; Length 14;
Best Local Similarity 61.5%; Pred. No. 4.4e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 APAAEGAGAEVKR 15
Db 2 AAAAAGAGPEWVR 14

RESULT 11

US-10-029-386-27605
; Sequence 27605, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 27605
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL139224.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
US-10-029-386-27605

Query Match 29.6%; Score 34; DB 14; Length 20;
Best Local Similarity 35.7%; Pred. No. 6.4e+02;
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TAPAAEGAGAEVKR 15
Db 7 SAPSVSLGCRIRR 20

RESULT 12

US-09-572-404B-1296
; Sequence 1296, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 1296
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in FKHL8 OR FREAC4 at 91-100 and may interact wi
; OTHER INFORMATION: Sequence 1295 in this patent.
US-09-572-404B-1296

Query Match 29.1%; Score 33.5; DB 10; Length 10;
Best Local Similarity 81.8%; Pred. No. 3.6e+02;
Matches 9; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 GTAPAAEGAGA 11
Db 1 GPAPAA-GAGA 10

RESULT 13

US-09-572-404B-3822
; Sequence 3822, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3822
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in FKHL8 OR FREAC4 at 91-100 and may interact wi
; OTHER INFORMATION: Sequence 3823 in this patent.
US-09-572-404B-3822

Query Match 29.1%; Score 33.5; DB 10; Length 10;
Best Local Similarity 81.8%; Pred. No. 3.6e+02;
Matches 9; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 GTAPAAEGAGA 11
Db 1 GPAPAA-GAGA 10

RESULT 14

US-08-344-824-276
; Sequence 276, Application US/08344824
; Publication No. US20030152580A1
; GENERAL INFORMATION:
; APPLICANT: SETTE, Alessandro
; APPLICANT: SIDNEY, John
; TITLE OF INVENTION: HLA BINDING PEPTIDES AND THEIR USES
; NUMBER OF SEQUENCES: 399
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Street Tower, 20th
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/344,824
; FILING DATE: 23-NOV-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,634
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 14137-80-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 276:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-344-824-276

Query Match 28.7%; Score 33; DB 8; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.2e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAAEGAGA 11
|||:|
Db 2 PAAAGVGA 9

RESULT 15
US-09-824-438-16
; Sequence 16, Application US/09824438
; Publication No. US20030073621A1
; GENERAL INFORMATION:
; APPLICANT: Butler, Jon P.
; APPLICANT: Hale, John E.
; APPLICANT: Heath Jr., William F.
; APPLICANT: Schoner, Brigitte E.
; APPLICANT: Heiman, Mark L.
; APPLICANT: Becker, Gerald W.
; APPLICANT: Varshavsky, Alexander D.
; TITLE OF INVENTION: Beta-lipotropin and Uses Thereof
; FILE REFERENCE: X-12139
; CURRENT APPLICATION NUMBER: US/09/824,438
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-824-438-16

Query Match 28.7%; Score 33; DB 10; Length 14;
Best Local Similarity 66.7%; Pred. No. 6.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAAEGAGAE 12
||:|
Db 2 PADDGGAQ 10

Search completed: October 7, 2004, 18:10:48
Job time : 85.5 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:50:43 ; Search time 22 Seconds
(without alignments)
56.319 Million cell updates/sec

Title: US-10-802-644-2

Perfect score: 115

Sequence: 1 GTAPAAEGAGAEVVKRAAEAKQAF 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 189146

Minimum DB seq length: 0

Maximum DB seq length: 24

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
 - 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
 - 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
 - 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
 - 5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
 - 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	33.0	20	US-08-053-451B-171	Sequence 171, Appl
2	36	31.3	21	US-09-319-570A-4	Sequence 4, Appli
3	36	31.3	22	US-09-236-423-1	Sequence 1, Appli
4	36	31.3	23	US-07-841-591A-14	Sequence 14, Appl
5	36	31.3	23	PCT-US93-02034-14	Sequence 14, Appl
6	34	29.6	21	US-08-726-306A-15	Sequence 15, Appl
7	33	28.7	14	US-09-217-228-16	Sequence 16, Appl
8	33	28.7	17	US-09-082-358B-30	Sequence 30, Appl
9	32	27.8	19	PCT-US91-02942-67	Sequence 67, Appl
10	32	27.8	20	US-09-579-664B-27	Sequence 27, Appl
11	32	27.8	21	US-08-557-309B-58	Sequence 58, Appl
12	32	27.8	21	US-08-929-414-11	Sequence 11, Appl
13	32	27.8	21	US-08-834-306-56	Sequence 56, Appl
14	32	27.8	21	US-08-834-306-58	Sequence 58, Appl
15	32	27.8	21	US-08-993-674A-56	Sequence 56, Appl
16	32	27.8	21	US-08-993-674A-58	Sequence 58, Appl
17	32	27.8	21	US-09-256-976-56	Sequence 56, Appl
18	32	27.8	21	US-09-256-976-58	Sequence 58, Appl
19	32	27.8	22	US-07-916-034-22	Sequence 22, Appl
20	31.5	27.4	21	US-08-557-309B-60	Sequence 60, Appl
21	31.5	27.4	21	US-08-929-414-12	Sequence 12, Appl
22	31.5	27.4	21	US-08-834-306-59	Sequence 59, Appl
23	31.5	27.4	21	US-08-993-674A-59	Sequence 59, Appl
24	31.5	27.4	21	US-09-256-976-59	Sequence 59, Appl
25	31	27.0	10	5240848-2	Patent No. 5240848
26	31	27.0	12	US-08-742-243-4	Sequence 4, Appli
27	31	27.0	14	5219739-39	Patent No. 5219739

28	31	27.0	16	US-08-660-592-11	Sequence 11, Appl
29	31	27.0	16	US-08-437-013-16	Sequence 16, Appl
30	31	27.0	16	US-09-275-508A-16	Sequence 16, Appl
31	31	27.0	17	US-09-244-583-10	Sequence 10, Appl
32	31	27.0	19	US-08-660-592-10	Sequence 10, Appl
33	30.5	26.5	20	US-09-247-527-18	Sequence 18, Appl
34	30	26.1	12	US-08-433-854-25	Sequence 25, Appl
35	30	26.1	12	US-08-174-745A-25	Sequence 25, Appl
36	30	26.1	12	US-08-195-947-25	Sequence 25, Appl
37	30	26.1	12	US-08-433-885-25	Sequence 25, Appl
38	30	26.1	12	US-08-433-908B-25	Sequence 25, Appl
39	30	26.1	12	US-08-410-614-25	Sequence 25, Appl
40	30	26.1	16	US-08-333-565-26	Sequence 26, Appl
41	30	26.1	16	US-07-942-245-105	Sequence 105, App
42	30	26.1	16	US-08-661-473-26	Sequence 26, Appl
43	30	26.1	18	US-08-105-454-6	Sequence 6, Appli
44	30	26.1	18	US-09-822-624-17	Sequence 17, Appl
45	30	26.1	19	US-09-050-739-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-08-053-451B-171
; Sequence 171, Application US/08053451B
; Patent No. 5955584
; GENERAL INFORMATION:
; APPLICANT: Chen, Francis W.
; APPLICANT: Ditlow, Charles C.
; APPLICANT: Calenoff, Emanuel
; TITLE OF INVENTION: ATHEROSCLEROTIC PLAQUE SPECIFIC
; TITLE OF INVENTION: ANTIGENS, ANTIBODIES THERETO, AND USES THEREOF
; NUMBER OF SEQUENCES: 176
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/053,451B
; FILING DATE: 26-APR-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: 7606-033-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-3660
; TELEFAX: 415-854-3694
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 171:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-053-451B-171

Query Match 33.0%; Score 38; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 10; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 1 GTAPAAEGAGA--EVKRAA 18
| : | ||| | : ||| |
| : | ||| | : ||| |

Db 1 GESPLTFGAGTKLEKRA20

RESULT 2

US-09-319-570A-4

Sequence 4, Application US/09319570A

Patent No. 6537760

GENERAL INFORMATION:

APPLICANT: BERGMANN, ANDREAS

APPLICANT: STRUCK, JOACHIM

TITLE OF INVENTION: RECEPTOR BINDING ASSAY FOR DETECTING TSH RECEPTOR

TITLE OF INVENTION: AUTOANTIBODIES AND REAGENT KIT FOR CARRYING OUT SUCH A

TITLE OF INVENTION: RECEPTOR BINDING ASSAY

FILE REFERENCE: 011377-0260244

CURRENT APPLICATION NUMBER: US/09/319,570A

CURRENT FILING DATE: 1999-08-04

PRIOR APPLICATION NUMBER: DE 196 51 093.7

PRIOR FILING DATE: 1996-12-09

NUMBER OF SEQ ID NOS: 5

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 4

LENGTH: 21

TYPE: PRT

ORGANISM: Homo sapiens

US-09-319-570A-4

Query Match 31.3%; Score 36; DB 4; Length 21;

Best Local Similarity 47.1%; Pred. No. 64;

Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 8 GAGAEVVKRASAEAKQAF 24

Db 5 GFGQELKNPQETLQAF 21

RESULT 3

US-09-236-423-1

Sequence 1, Application US/09236423

Patent No. 6395875

GENERAL INFORMATION:

APPLICANT: Freimuth, Paul

TITLE OF INVENTION: RECOMBINANT SOLUBLE ADENOVIRUS RECEPTOR

FILE REFERENCE: BNL-2007

CURRENT APPLICATION NUMBER: US/09/236,423

CURRENT FILING DATE: 1999-01-25

NUMBER OF SEQ ID NOS: 9

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1

LENGTH: 22

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: CLONING VECTOR

US-09-236-423-1

Query Match 31.3%; Score 36; DB 4; Length 22;

Best Local Similarity 47.4%; Pred. No. 68;

Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 PAAGAGAEVVKRASAEAKQ 22

Db 4 PAANKARKEAEIAAATAEQ 22

RESULT 4

US-07-841-591A-14

Sequence 14, Application US/07841591A

Patent No. 5900476

GENERAL INFORMATION:

APPLICANT: Ruggeri, Zaverio M. and

APPLICANT: Ware, Jerry, inventors

APPLICANT: on behalf of The Scripps Research

US-10-802-644-2.closed.ra1

APPLICANT: Institute

TITLE OF INVENTION: Therapeutic Domains of

TITLE OF INVENTION: von Willebrand Factor

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: The Scripps Research Institute

STREET: 10666 No. 5900476th Torrey Pines Road

CITY: La Jolla

STATE: California

COUNTRY: United States

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: 1.2 megabyte 5 1/4" floppy

COMPUTER: AST Bravo 386SX

OPERATING SYSTEM: MS DOS version 3.2

SOFTWARE: WordPerfect 5.1 conv. to ASCII

CURRENT APPLICATION DATA: US/07/841,591A

APPLICATION NUMBER: US/07/841,591A

FILING DATE: 26-Feb-92

CLASSIFICATION:

PRIOR APPLICATION DATA: This is a c-i-p of

APPLICATION NUMBER: PCT/US91/07756

FILING DATE: 17-Oct-91

ATTORNEY/AGENT INFORMATION:

NAME: Barron, Alexis

REGISTRATION NUMBER: 22,702

REFERENCE/DOCKET NUMBER: P16,633-H

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 923-4466

TELEFAX: (215) 923-2189

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 23

TYPE: Amino acid

STRANDEDNESS:

TOPOLOGY: Linear

US-07-841-591A-14

Query Match 31.3%; Score 36; DB 2; Length 23;

Best Local Similarity 47.4%; Pred. No. 71;

Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 PAAGAGAEVVKRASAEAKQ 22

Db 5 PAANKARKEAEIAAATAEQ 23

RESULT 5

PCT-US93-02034-14

Sequence 14, Application PC/TUS9302034

GENERAL INFORMATION:

APPLICANT: Ruggeri, Zaverio M. and

APPLICANT: Ware, Jerry, inventors

APPLICANT: on behalf of The Scripps Research

APPLICANT: Institute

TITLE OF INVENTION: Therapeutic Domains of

TITLE OF INVENTION: von Willebrand Factor

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: The Scripps Research Institute

STREET: 10666 North Torrey Pines Road

CITY: La Jolla

STATE: California

COUNTRY: United States

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: 2 megabyte 3.25" floppy

COMPUTER: AST Bravo 386SX

OPERATING SYSTEM: MS DOS version 3.2

SOFTWARE: WordPerfect 5.1 conv. to ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/02034

FILING DATE: 19930223

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA: This is a c-i-p of
; APPLICATION NUMBER: PCT/US91/07756
; FILING DATE: 17-Oct-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Barron, Alexis
; REGISTRATION NUMBER: 22,702
; REFERENCE/DOCKET NUMBER: P16,633-H PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 923-4466
; TELEFAX: (215) 923-2189
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: Linear
; PCT-US93-02034-14

Query Match 31.3%; Score 36; DB 5; Length 23;
Best Local Similarity 47.4%; Pred. No. 71;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 PAAGAGAEVVKRAAEAKQ 22
   ||| | | : | | | |
Db 5 PAANKARKEAEIAATAAEQ 23

RESULT 6
US-08-726-306A-15
; Sequence 15, Application US/08726306A
; Patent No. 5958684
; GENERAL INFORMATION:
; APPLICANT: van Leeuwen, Frederik Willem
; APPLICANT: Burbach, Johannes Peter Henri
; APPLICANT: Grosveld, Franklin G.
; TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
; NUMBER OF SEQUENCES: 189
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1 Financial Center
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,306A
; FILING DATE: 02-Oct-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 95/20080.4
; FILING DATE: 02-Oct-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,832
; FILING DATE: 01-Jan-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Ph.D., Kathleen M.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 345-9100
; TELEFAX: (617) 345-9111
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide

```

```

US-08-726-306A-15
Query Match 29.6%; Score 34; DB 2; Length 21;
Best Local Similarity 61.5%; Pred. No. 1.3e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 1 GTAPAA--EGAGA 11
   ||||| : |||
Db 5 GTAPAGRLDGSQA 17

RESULT 7
US-09-217-228-16
; Sequence 16, Application US/09217228
; Patent No. 6323178
; GENERAL INFORMATION:
; APPLICANT: Butler, Jon P.
; APPLICANT: Hale, John E.
; APPLICANT: Heath Jr., William F.
; APPLICANT: Schoner, Brigitte E.
; APPLICANT: Heiman, Mark L.
; APPLICANT: Becker, Gerald W.
; APPLICANT: Varshavsky, Alexander D.
; TITLE OF INVENTION: Beta-lipotropin and Uses Thereof
; FILE REFERENCE: X-12139
; CURRENT APPLICATION NUMBER: US/09/217,228
; CURRENT FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-217-228-16

Query Match 28.7%; Score 33; DB 4; Length 14;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAAGAGAE 12
   || : |||
Db 2 PADDGAGAQ 10

RESULT 8
US-09-082-358B-30
; Sequence 30, Application US/09082358B
; Patent No. 6469153
; GENERAL INFORMATION:
; APPLICANT: Goff, Stephen P.
; APPLICANT: Li, Xinguiang
; TITLE OF INVENTION: EIP-1, EIP-3 GENES, ENVELOPE-INTERACTING PROTEINS,
; FILE REFERENCE: EIP-1, and EIP-3
; FILE REFERENCE: 0575/54804
; CURRENT APPLICATION NUMBER: US/09/082,358B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 17
; TYPE: PRT
; ORGANISM: murine
US-09-082-358B-30

Query Match 28.7%; Score 33; DB 4; Length 17;
Best Local Similarity 70.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 EGAGAEVVKRA 16
   : ||||| | |
Db 1 DGAGARVLRA 10

```

RESULT 9

PCT-US91-02942-67
; Sequence 67, Application PC/TUS9102942
; GENERAL INFORMATION:
; APPLICANT: ROTHLEIN, ROBERT
; APPLICANT: ADARI, JOHN R
; APPLICANT: ATHWAL, DILJEET S
; TITLE OF INVENTION: HUMANIZED CDR-GRAFTED ICAM-1 ANTIBODY
; NUMBER OF SEQUENCES: 102
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Ave. NW Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/02942
; FILING DATE: 19910429
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9009549.8
; FILING DATE: 27-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: FOX, SAM L
; REGISTRATION NUMBER: 30,353
; REFERENCE/DOCKET NUMBER: 1011.0586600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 466-0800
; TELEFAX: (202) 833-8716
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US91-02942-67

Query Match 27.8%; Score 32; DB 5; Length 19;
Best Local Similarity 46.2%; Pred. No. 2.2e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 9 AGAEVKRASAEAK 21
:|||||: :
DB 2 SGAEVKPGSSVK 14

RESULT 10

US-09-579-664B-27
; Sequence 27, Application US/09579664B
; Patent No. 6514719
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
; FILE REFERENCE: 2923-A
; CURRENT APPLICATION NUMBER: US/09/579,664B
; CURRENT FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 27
; LENGTH: 20
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: peptide
US-09-579-664B-27

Query Match 27.8%; Score 32; DB 4; Length 20;
Best Local Similarity 47.4%; Pred. No. 2.3e+02;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 5 AAGAGAEVKRASAEAKQA 23
:|||||: :
DB 2 AAAREPAEPSPAAAEAGA 20

RESULT 11

US-08-557-309B-58
; Sequence 58, Application US/08557309B
; Patent No. 5916572
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,309B
; FILING DATE: 14-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.422
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-5031
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-557-309B-58

Query Match 27.8%; Score 32; DB 2; Length 21;
Best Local Similarity 47.1%; Pred. No. 2.4e+02;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 7 EGAGAEVKRASAEAKQA 23
:|||||: :
DB 1 KAAAPAKAAAPAKAA 17

RESULT 12

US-08-929-414-11
; Sequence 11, Application US/08929414
; Patent No. 5942403
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Houghton, Raymond
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION
; OF T. CRUZI INFECTION

NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/929,414
FILING DATE: 15-SEP-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.406C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PS-08-929-414-11

```

; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-834-306-56

Query Match 27.8%; Score 32; DB 3; Length 21;
Best Local Similarity 47.4%; Pred. No. 2.4e+02;
Matches 8; Conservative 2; Mismatches 7; Indels

Qy 7 EGAGAEVVKRASAEAKQA 23
      :| | | | | | | | | |
Db 1 KAAAPAKAAAAAPAKQA 17
      :| | | | | | | | | |

RESULT 14
US-08-834-306-58
; Sequence 58, Application US/08834306
; Patent No. 6054135
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION
; OF SEQUENCES: 65
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; City: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,306
; FILING DATE: 15-APR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.422C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-834-306-58

```

RESULT 13
US-08-834-306-56
; Sequence 56, Application US/08834306
; Patent No. 6054135
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,306
; FILING DATE: 15-APR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.422C1
; TELECOMMUNICATION INFORMATION:

Patent No. 6228372
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond L.
APPLICANT: Smith, John M.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,674A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.422C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-993-674A-56
Query Match 27.8%; Score 32; DB 3; Length 21;
Best Local Similarity 47.1%; Pred. NO. 2.4e+02;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Oy 7 EGAGAEVVKASAEAKQA 23
Db 1 KAAAAAPAKAAAAAPAKAA 17

Search completed: October 7, 2004, 17:58:14
Job time : 23 secs